LEVEL II

(12)

AD A105072

AFGL-TR-81-0030

MILITARY GEODESY AND GEOSPACE SCIENCE
Unit Three

Warren G. Heller
A. Richard LeSchack

The Analytic Sciences Corporation
One Jacob Way
Reading, Massachusetts 01867

February 1981

Scientific Report No. 7

Approved for public release; distribution unlimited

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFGL-TR-81-0030	2. GOVT ACCESSION NO.: AD-A105 072
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): MILITARY GEODESY AND GEOSPACE SCIENCE. Unit Three.	5. TYPE OF REPORT & PERIOD COVERED: Scientific Report No. 7
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Warren G. Heller, A. Richard LeSchack	8. CONTRACT OR GRANT NUMBER(s): F19628-77-C-0152
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Analytic Sciences Corporation, One Jacob Way, Reading, Massachusetts 01867	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 320432AA
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Geophysics Laboratory, Hanscom AFB, Massachusetts 01731, Monitor/Brian Mertz, Capt, USAF/LWG	12. REPORT DATE: February 1981
	13. NUMBER OF PAGES: 172
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Geodesy, gravity, physical geodesy, mapping, charting

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This lecture course provides a full-year introduction to Military Geodesy and Geospace Science. Throughout the presentation a military perspective is maintained which links Mapping, Charting, and Geodesy (MC&G) issues with modern defense requirements. Elementary preparation is assumed in the subjects of general physics, mechanics, chemistry, astronautics, and linear system theory. The student should also be familiar with differential equations, analytic geometry, and linear algebra. Some acquaintance with vector calculus is useful but not essential.

DD FORM 1 JAN 73 1473

404565

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

The topics covered herein are intended to provide conceptual rather than working knowledge. Ideally, the student completing this course will have attained a broad understanding of the MC&G field and will be able to develop specialized expertise quickly when required.

The notes are intended to be presented in chapter/section order within each of the four Units of Instruction. However, several of the subsections in these notes contain more advanced material which may be omitted without loss of continuity. These subsections are denoted with the symbol (†) after the title. A fifth volume contains faculty material.

The organizational flow of the lectures is from concepts in the initial sections, particularly in Unit One, to applications and specific systems later on. As a result the student is often referred ahead to provide motivation in regard to relevancy. In later chapters, however, the situation is reversed with the student referred back to review important conceptual material as necessary.

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

FOREWORD

This lecture course provides a full-year introduction to Military Geodesy and Geospace Science. Throughout the presentation a military perspective is maintained which links Mapping, Charting, and Geodesy (MC&G) issues with modern defense requirements. Elementary preparation is assumed in the subjects of general physics, mechanics, chemistry, astronautics, and linear system theory. The student should also be familiar with differential equations, analytic geometry, and linear algebra. Some acquaintance with vector calculus is useful but not essential.

The topics covered herein are intended to provide conceptual rather than working knowledge. Ideally, the student completing this course will have attained a broad understanding of the MC&G field and will be able to develop specialized expertise quickly when required.

The notes are intended to be presented in chapter/section order within each of the four Units of Instruction. However, several of the subsections in these notes contain more advanced material which may be omitted without loss of continuity. These subsections are denoted with the symbol (†) after the title.

The organizational flow of the lectures is from concepts in the initial sections, particularly in Unit One, to applications and specific systems later on. As a result the student is often referred ahead to provide motivation in regard to relevancy. In later chapters, however, the situation is reversed with the student referred back to review important conceptual material as necessary.

GLOSSARY OF ACRONYMS AND ABBREVIATIONS FOR UNIT THREE

BLIP	Background Limited Performance	3-127
CCD	Charge Coupled Device	3-129
CCTV	Closed-Circuit Television	3-95
CRT	Cathode Ray Tube	3-57
CW	Continuous Wave	3-119
EIO	Extended Interaction Oscillator	3-145
EM	Electromagnetic	3-123
ERTS	Earth Resources Technology Satellite (original designation for LANDSAT)	3-7
FLIR	Forward Looking Infrared	3-125
H and D Curve	Hurter and Driffield Curve (characteristic curve for photographic materials)	3-26
IIR	Imaging Infrared	3-123
IR	Infrared	3-133
KC	Kalvar Corporation (designation for a non-silver vesicular photographic process)	3-65
LANDSAT	A series of NASA satellites for viewing the earth at various frequencies	3-2
MMW	Millimeter Wave	3-138
MTF	Modulation Transfer Function	3-54
MW	Microwave	3-138
NM	Nanometer	3-10
OTF	Optical Transfer Function	3-81
PPDB	Point Positioning Data Base	3-150

GLOSSARY OF ACRONYMS AND ABBREVIATIONS FOR UNIT FOUR (Continued)

PSF	Point Spread Function	3-77
SAR	Synthetic Aperture Radar	3-147
SI	System International (the metric system of units based on meter/ kilogram/second	3-13
SKYLAB	A manned satellite program	3-6
SLAR	Side-Looking Airborne Radar	3-4
TDI	Time Delay and Integration	3-130
TIROS	Television Infrared Observation Satellite (NASA)	3-6
TGS	Triglycine Sulphate	3-128
UV	Ultraviolet	3-16

UNIT THREE

TABLE OF CONTENTS

UNIT THREE

TABLE OF CONTENTS (Continued)

UNIT THREE

LIST OF FIGURES

UNIT THREE

LIST OF FIGURES (Continued)

UNIT THREE

LIST OF FIGURES (Continued)

UNIT THREE

LIST OF TABLES

UNIT THREE
SOURCE DATA COLLECTION AND REMOTE SENSING

CHAPTER ONE
INTRODUCTION

The object of Unit Three is to introduce the reader to a variety of material related to the collection and evaluation of data by the use of sensing devices located at great distances from the objects being investigated. The general concept is known as remote sensing and is defined in more detail in Section 3.1.1, along with a brief review of its historical development (Section 3.1.2) and prospects for future growth (Section 3.1.3).

The major areas included within Unit Three are imaging systems (Chapter Two) and imaging operations (Chapter Three). The term image is used in a general sense to include not only a conventional picture of a distant object, but also the results obtained from measuring systems involving forms of energy other than visible light -- such as, for example, infrared (Chapter Five) or microwave (Chapter Six) energy.

The main concepts to be discussed under the heading of imaging systems are:

- Photographic systems
- Electro-optical imaging systems
- Digital image characteristics

- Non-silver imaging materials
- Electrostatic and vesicular film systems
- Solid-state focal plane imaging systems
- The modulation transfer function.

The concept of imaging operations refers to the modification or transformation of the data collected in order to extract as much information as possible or to produce an image more suitable for viewing, analysis, and interpretation. Many of these operations are carried out by computer, and require that the image be scanned and digitized for input to the computer.

Chapter Three begins by discussing geometric transformations of images, as well as the processes of scanning and digitizing. The other major topic under imaging operations is image enhancement, which includes a wide variety of techniques for improving the visual information content of a picture and for detecting features of interest that may be difficult or impossible to discern in the original image. An extended discussion of a particular remote-sensing system called LANDSAT is reserved for Unit Four.

3.1.1 Definition of Remote Sensing

Remote sensing refers to techniques used to obtain information about objects, areas, or phenomena by analyzing data collected by devices located at a distance from the object of investigation. This is a very broad definition, which could be construed to include, for example, the process of reading -- in which the eye (as the sensing device) obtains information by analyzing patterns of high reflectivity and low reflectivity on the printed page. For these patterns to have

any meaning, they must be analyzed and interpreted -- a process which, in the case of reading, takes place in the retina, the optic nerve, and the brain. Practical remote sensing systems make use of the phenomena of emission, absorption, and reflection of various kinds of energy. Sonar systems (not discussed further) are an example of remote sensing using sound or acoustical energy. The remote sensing systems considered in Unit Three use various forms of electromagnetic energy which may be emitted by the object under study or reflected from it. In particular, the type of energy may be visible light, ultraviolet light, infrared energy, or microwave (radar) energy. In most applications of inte st, the objects or phenomena to be studied are located on or near the ground or surface, while the remote sensing device is in an aircraft or a satellite. A general overview of the concept of remote sensing is shown in Figure 3.1-1.

Remote sensing involving light, ultraviolet, or infrared energy generally makes use of techniques related to photography or television (electro-optical imaging). For this reason, considerable space is devoted to these topics in Chapter Two. Non-photographic techniques for the detection of infrared energy are also important, and are usually referred to as thermography. Microwave energy detection may be classified as:

- Active - in which energy produced by the remote sensing device is reflected or scattered by the object under investigation and the returning energy is detected by the device (radar)
- Passive - in which the object under investigation emits microwave energy or reflects microwave energy from the sun or other sources, and the remote sensing radiometer receives the emitted or reflected energy.

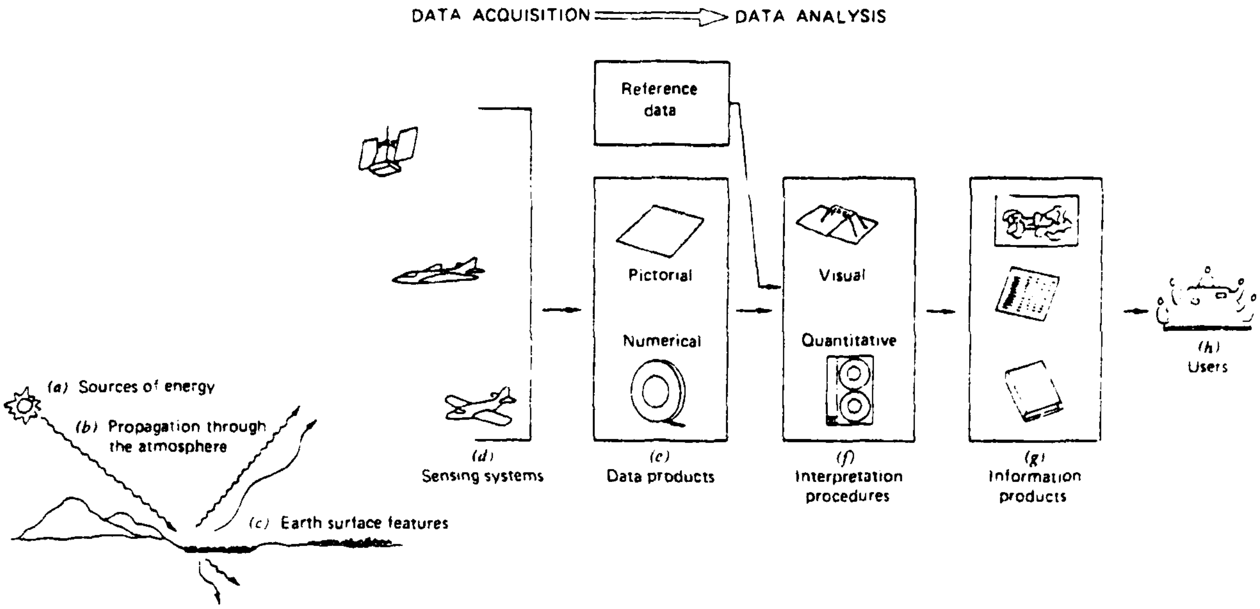
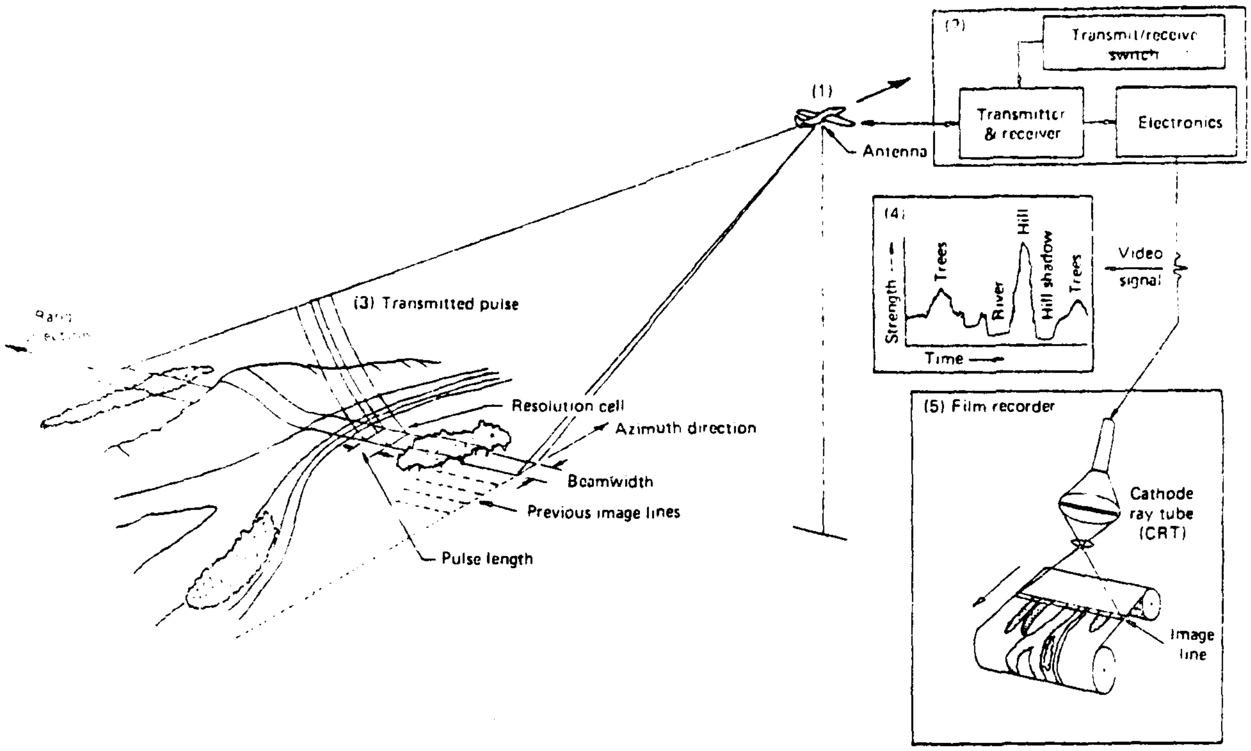


Figure 3.1-1 Concepts of Remote Sensing

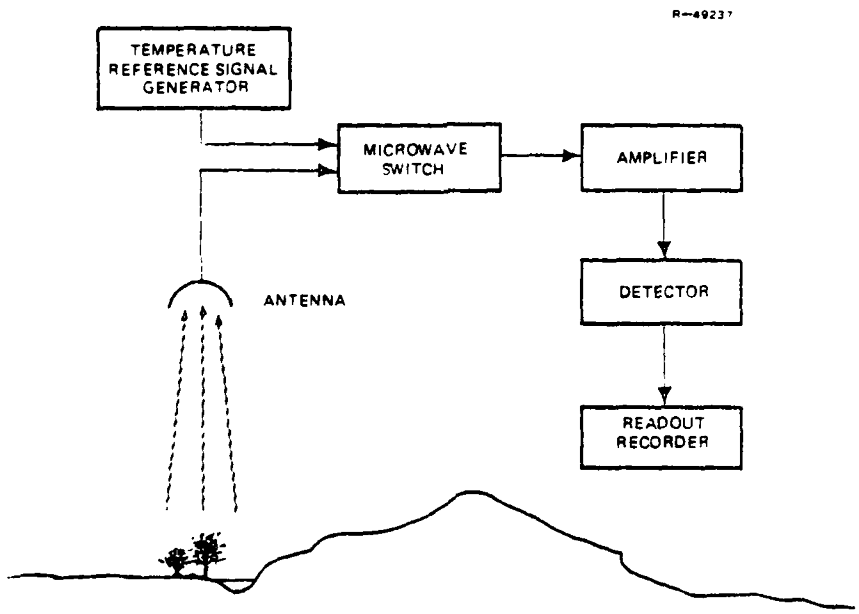
An important example of a radar sensing system is side-looking airborne radar (SLAR). Figure 3.1-2 illustrates the basic principles of SLAR (active microwave sensing), as well as passive microwave sensing.

3.1.2 Historical Perspective

It may surprise the reader to learn that the first known aerial photograph was obtained in 1858, using a balloon as the remote platform. Kites were also used as platforms for aerial photography as early as 1882. A photographic survey of the effects of the San Francisco earthquake and fire of 1906 was made by the use of a specially constructed camera elevated to an altitude of approximately 600 m using an array of kites. The camera produced enormous negatives (1.4 by 2.4 m), showing considerable detail in the image. The first recorded photography



A. Active Microwave Sensing (Side-looking Airborne Radar)



B. Passive Microwave Sensing

Figure 3.1-2 Microwave Sensing

from an airplane took place in 1906 when a motion picture photographer accompanied Wilbur Wright on a training flight in Italy. By the time of the first World War aerial photography was a recognized tool of military reconnaissance.

Remote sensing from space also has a long history. Concepts of rocket-propelled camera systems go back to the beginning of this century. In 1912, a large-format camera was boosted to a height of 800 m by an experimental rocket. Actual remote sensing from space began in 1946 when small cameras were carried aboard captured V2 rockets fired from the White Sands Proving Grounds in New Mexico. Early meteorological satellites (TIROS-1, 1960, for example) returned views of cloud patterns and occasional images of the earth's surface. Photography with hand-held cameras was part of the Mercury, Gemini, and Apollo missions of the 1960s. Results of these missions not only proved the validity of the concept of photographic remote sensing from space, but also led to significant discoveries in the earth and environmental sciences. During the Apollo 9 earth orbit flight, the concept of multispectral photography for earth resources studies was demonstrated. The experiment used a battery of four cameras, simultaneously photographing the same scene with different combinations of film-type and filters, to obtain views in different regions of the visible and infrared spectrum. Skylab astronauts obtained more than 35,000 photographic images of the earth with a variety of remote sensing devices, including a 6-camera multispectral array, a long-focal-length camera for high resolution photographs, a 13-channel multispectral scanner, and two microwave systems.

A series of satellites specifically designed to acquire data on earth resources by remote sensing was planned in 1967, under a program designated as Earth Resources Technology

Satellites (ERTS). The first successful vehicle, known as ERTS-1, operated from 1972 to 1978. Later satellites in this program were redesignated LANDSAT. The LANDSAT system is considered in some detail in Unit Four.

3.1.3 Future Growth

There are major trends in the development and application of remote-sensing systems that can be extrapolated into the future. Remarkable improvements in the quality and sensitivity of imaging devices, coupled with advances in the technology of image processing and enhancement, can be expected to continue. The amount of information and level of detail that will be available by the use of remote sensing systems on board satellites may stagger the imagination. Other trends may decrease the cost of placing satellites into near-earth orbit. For example, the Space Shuttle and its successors may well be used for the placement and recovery of orbiting remote-sensing platforms of all kinds. Remote sensing from space is destined to play a growing role in a variety of military as well as non-military applications.

CHAPTER TWO

IMAGING SYSTEMS

The formation of images involves two major technologies:

- Photochemical (e.g., conventional photography) based on the chemical effects of light energy

- Photo-electronic or electro-optical (e.g., television) based on the electronic effects of light energy.

Both technologies are reviewed at some length in this chapter. The discussion begins with photography, a general name given to processes in which pictures are produced by the action of light. Such pictures could, in principle, be produced by any material that is sensitive to light. There are many such materials. Some fade under the action of light; others darken; some change in various properties because their molecules are decomposed by light; others change because their molecules condense and link up to form long chains or other structures (e.g., polymerization); some are rendered more soluble by the action of light; others, less soluble. Any one of these changes could, in principle, be used as the basis of a photographic process.

It is because of their <u>extreme sensitivity</u> to light (in comparison with other photochemical processes) that silver compounds are used for most types of photography. When such compounds decompose under the influence of light, metallic silver is formed in a finely divided state. The metallic silver particles show none of the bright metallic appearance characteristic of large masses of silver. They are, in fact, black

in color, thus accounting for the characteristic darkening of silver-based films and papers.

Photo-electronic imaging devices are discussed next. Such devices can be competitive with photographic systems in a number of applications; each technology has its advantages, as well as disadvantages. From an overall point of view, some of these comparisons are summarized in Table 3.2-1.

Subsequent topics will be:

- Digital imagery characteristics
- Non-silver imaging materials
- Electrostatic and vesicular film (KC) systems
- Solid-state focal plane imaging systems
- Modulation transfer function.

TABLE 3.2-1

COMPARISON OF PHOTOGRAPHIC AND PHOTO-ELECTRONIC SYSTEMS

SYSTEM	ADVANTAGES	DISADVANTAGES
Photographic	Superior resolution; convenience of storage and access	Requirement for chemical processing; environmental sensitivity of stored material (particularly prior to processing)
Photo-electronic	Direct access; image information in electronic form to facilitate transmission, storage, and computer processing of images	Limit to resolution; power requirements; equipment cost

Before proceeding with the study of individual imaging systems, it is appropriate to review some basic facts concerning light and how it is measured.

Light is regarded, in physics, as an example of energy propagated by electromagnetic waves; it differs from non-visible radiation only in its wavelength. Visible light consists of energy having wavelengths between 380 and 760 nm* as shown in Fig. 3.2-1. Those regions of nonvisible energy adjacent to the visible (i.e., ultraviolet and infrared) are also important to the material of this chapter. Light is also characterized, in its interactions with matter, by its quantum nature. Light energy is emitted or absorbed only in discrete units called quanta or photons. The energy of each photon depends on the wavelength (color) of the light, but not on its intensity. Intensity is proportional to the number of photons. Light energy is thus considered in a dual fashion: as electromagnetic waves characterized by wavelength and intensity or as a stream of photons, the energy of which depends on the wavelength.

The measurement of light is a subject that is fundamental to the study of photography and other imaging systems. When thought of as a form of electromagnetic energy, light is measured in the ordinary units of physics -- for example, the rate at which energy is being emitted is expressed in units of joules per second or watts. Measurements of this type are called radiometric, and have the disadvantage that they do not take into account the variable sensitivity of the human eye to light of different wavelengths. Two sources of different color, emitting the same number of watts of energy, would not appear equally bright to an observer. For this reason there is a second system of measurement, known as photometric, which is

*One nanometer, abbreviated nm, equals 10^{-9} meters.

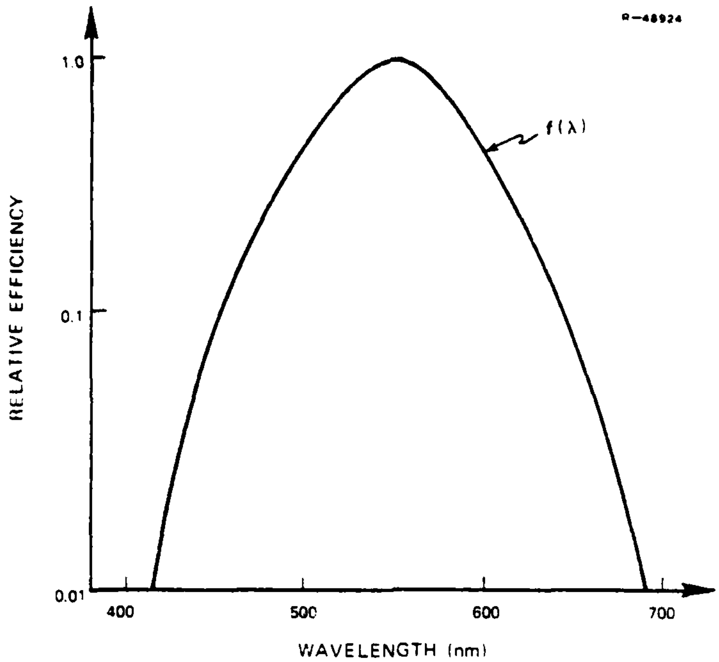


Figure 3.2-2 Relative Spectral Efficiency of the Human Eye

where

Φ_P is the observed photometric effect, in lumens

$\vartheta_R(\lambda)$ is the distribution of energy by wavelength (watts/meter)

$f(\lambda)$ is the relative efficiency curve of Fig. 3.2-2

K is a constant for the adjustment of units from the radiometric to the photometric scale

The integral is taken over the range of wavelengths for which $f(\lambda)$ is significantly different from zero -- that is, the visible range (denoted as λ_1 and λ_2 in Eq. 3.2-1). Actual lumen-to-watt ratios range from about 150 for high-efficiency mercury or sodium vapor lamps (used for outdoor illumination), down to

PRECEDING PAGE BLANK-NOT FILMED

15 for ordinary incandescent lights (because so much of the energy is in the infrared range).

Table 3.2-2 shows the correspondence between the most important quantities in the radiometric and photometric systems. Irradiance (and illuminance) measure the energy falling on a given surface; they formalize the concept of intensity of illumination. Radiant intensity (and luminous intensity) characterizes the apparent strength of a source of energy in a particular direction, by measuring power emitted into a unit solid angle.* Radiance (and luminance) describe the apparent surface brightness of an extended energy emitter in a given direction. The reader with experience in the use of photographic light meters will recognize that:

- Incident light meters measure illuminance
- Spot meters measure luminance of individual objects in the scene
- Reflected light meters, including most built-in meters in automatic or semi-automatic cameras, measure the average luminance of the various objects included in the field of view.

While the photometric units given in Table 3.2-2 are the standard scientific units (SI units) now adopted internationally for scientific and engineering use, the older units, based on the English system, are still widely used in the United States. The correspondences and conversions are given in Table 3.2-3.

---

*The unit of solid angle is the steradian. Solid angles are measured by the area subtended on the surface of a unit sphere centered at the apex of the solid angle. For a solid angle of one steradian, the subtended area is one. There are 4π steradians in the entire sphere.

TABLE 3.2-2

RADIOMETRIC AND PHOTOMETRIC QUANTITIES

RADIOMETRIC		PHOTOMETRIC	
NAME	UNITS	NAME	UNITS
Radiant flux	Watt	Luminous flux	Lumen
Irradiance	Watt/meter^2	Illuminance	Lumen/meter^2 (called lux)
Radiant intensity	Watt/steradian	Luminous intensity (formerly called candlepower)	Lumen/steradian, (called candela)
Radiance	Watt/(steradian · meter^2)	Luminance	Candela/meter^2 (called nit)

TABLE 3.2-3

CORRESPONDENCE BETWEEN ENGLISH AND METRIC (SI) PHOTOMETRIC UNITS

QUANTITY	SI UNIT	ENGLISH UNIT	CONVERSION
Illuminance	Lux	Footcandle	1 Footcandle = 10.76 Lux
Luminance	Nit	Foot lambert	1 Foot lambert = 3.43 Nit
		Candle/foot^2	1 Candle/foot^2 = 10.76 Nit

NOTE: 1 Foot lambert = 0.32 Candle/foot^2

3.2.1 Photographic Systems

Conventional photographic systems are based on the extreme sensitivity to light of a class of silver compounds called the silver halides -- silver chloride, silver bromide,

and silver iodide. (Silver fluoride also belongs to this class of compounds, but is not usually included because its high solubility in water makes it unsuitable for most photographic applications.) Because of the key role that photographic technology plays in reconnaissance and remote sensing, it is important for the student to have an appreciation of the capabilities and limitations of the various types of photographic materials. For this reason, some time will now be devoted to the study of basic photographic theory and principles. The material to be covered includes:

- Properties and preparation of light-sensitive materials
- The nature of the photographic latent image
- Development and processing of photographic materials
- Properties of film and images
- Materials and processes for color photography
- Materials and processes for instant photography
- Properties and preparation of light-sensitive materials.

The basic light-sensitive material in conventional photographic systems is a uniform dispersion of microcrystals of silver halide in a water-permeable protective colloid, usually gelatin. This combination is called a photographic emulsion. The gelatin serves to prevent aggregation of the silver halide microcrystals, thus maintaining optimum spacing and distribution, and helps to control the growth and crystal form of the halide particles. It also plays various roles, to some extent not entirely understood, in image formation and processing. For example, minute traces of deliberately introduced

impurities (like sulphur) may have a beneficial effect on the properties of the resulting photographic emulsion.

Photographic emulsions are generally prepared by allowing the gelatin to soak up a soluble silver salt such as silver nitrate, and then, in total darkness, soaking the material in a solution of a soluble halide salt such as potassium bromide, to form crystals of insoluble -- and light sensitive -- silver halide within the emulsion. A number of subsequent chemical and physical processes are required to achieve the proper size, form, and distribution of the silver halide microcrystals. These processes are usually referred to as ripening, setting, and digestion. Depending on the desired characteristics of the emulsion and the uses for which it is intended, the sizes of the microcrystals (or <u>grains</u>) may range from as little as 0.1 μm up to as much as 5 μm.

Most photographic materials use a mixture of two or three silver halides, rather than pure silver bromide, etc. Such mixed silver halide emulsions do not consist of separate crystals of silver bromide, silver chloride, and silver iodide. Instead, all the crystals are similar, and contain the halides in about the same proportion. Silver chloride is colorless and is sensitive only to the ultraviolet and extreme violet end of the visible spectrum. Silver bromide is pale yellow; silver iodide is somewhat deeper yellow. The sensitivity of silver bromide extends from the ultraviolet (UV) to 490 nm (blue). The addition of silver iodide deepens the color and extends the sensitivity to 520 nm (blue-green). The main function of the silver iodide component is to increase sensitivity to light, and also to contribute to the formation of more uniform and smaller grains.

Chemical treatments to increase the light sensitivity of the emulsion are referred to as sensitization. This is done by the carefully controlled addition of very small amounts of certain chemical compounds to the emulsion. Some examples of chemical sensitizing agents are:

- Compounds containing sulphur
- Compounds of gold and/or palladium.

These need be introduced only in extremely minute quantities to have the desired effect.

A very important step remains in the preparation of a usable photographic emulsion. Since silver halide crystals are sensitive only to UV, violet, and blue light, it is necessary, for most photographic purposes, to extend the sensitivity to cover more of the visible spectrum. There are four general categories of emulsions, depending on how far in the direction of longer wavelengths of light this sensitivity is extended:

- Natural, or blue-sensitive
- Orthochromatic, in which the sensitivity extends into the green
- Panchromatic, in which the sensitivity includes the entire visible range
- Infrared, in which the sensitivity extends through the visible region and into the infrared. If filters are used to block out visible light, such film responds only to infrared.

This process of extending the response of an emulsion is known as spectral sensitization. It is effected by introducing a small amount of a sensitizing dye into the emulsion, which can

absorb light in the longer wavelength region of the spectrum -- light to which the silver halides would normally be insensitive -- and, in some manner, transfer the energy thus absorbed to the molecules of silver halide. The actual mechanisms by which sensitizing dyes operate can not yet be fully explained, although a number of theories have been suggested. Whichever interpretation may be the correct one, the dye-sensitized crystals are affected by light in the same manner as if the light had been directly absorbed by the non-dyed crystal.

A number of other chemicals are usually added to the spectrally sensitized emulsion to complete the manufacturing process. These include:

- Stabilizers
- Wetting agents
- Hardening agents, to insure that the gelatin layer is tough enough to withstand handling
- Plasticizers, to insure that the emulsion remains flexible and does not crack
- Bacteriocides, to inhibit bacterial attack during storage.

The completed emulsion is applied in a uniform thin layer to a suitable backing material or base (sometimes also called the substrate). For photographic films, the base is a flexible plastic like cellulose acetate or polyester (mylar). Emulsions coated on thin sheets of glass are called photographic plates. Specially whitened, high-quality paper or thin cardboard, sometimes treated with a water-resistant surface coating to make it impervious to processing chemicals, is the base for photographic papers. Emulsions can also be coated, for special applications, on cloth, metal, and many other materials.

Exposure to light affects the photographic emulsion by causing subtle changes in the internal crystal structure of the halide grains, making the altered grains susceptible to chemical decomposition during the subsequent development process. The decomposition, or development, of a halide microcrystal leads to the formation of metallic silver. In dark areas of the image, where the halide microcrystals have not been affected by light, the developer has no effect and no metallic silver is formed. The distribution, within the emulsion, of altered and unaltered halide microcrystals, in accordance with the light intensities of the image, is called the latent image.

The mechanics of latent image formation, and the development process in which the latent image is made visible, are now discussed in detail.

The nature of the photographic latent image - The most widely accepted theory of the mechanism of photographic image formation is described below. A few preliminary concepts are reviewed before the theory is outlined.

When atoms of silver and one of the halogens (bromine, for example) form the regular lattice structure that makes up the crystal of silver bromide, the configuration of the electron charge distributions associated with the constituent atoms is altered by the presence of surrounding atoms. The outermost (valence) electron of the silver atom becomes associated with a nearby bromine atom, leading to a stable closed-shell configuration for the two atoms. The new configuration is electrically charged by the gain (or loss) of the exchanged electron (Fig. 3.2-3). The lattice structure of silver bromide, then, involves regularly spaced silver and bromine ions. Since all electrons are strongly bound, the

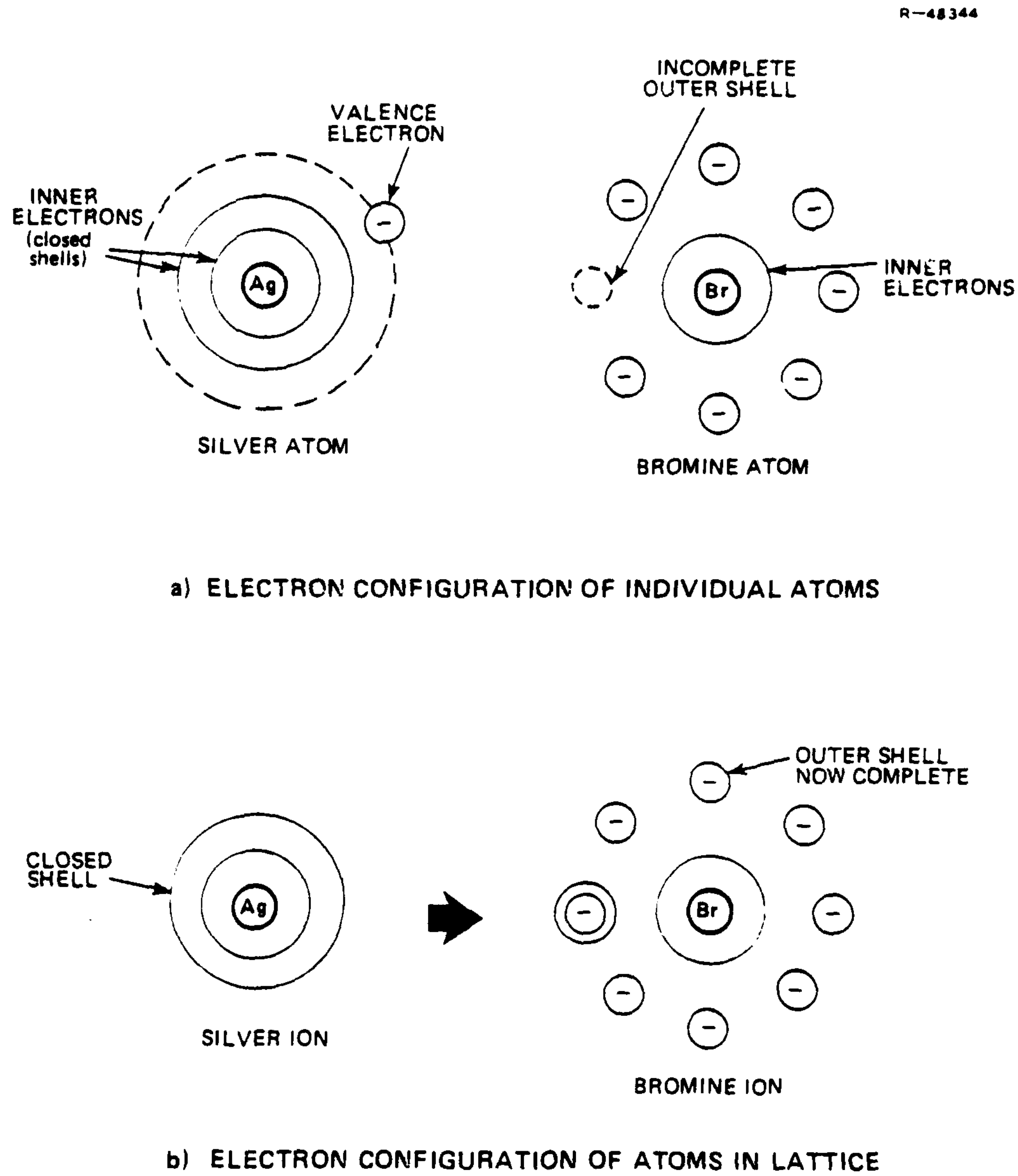


Figure 3.2-3 Silver Halide Lattice Structure

electrical conductivity of silver bromide is very low. Metallic silver, on the other hand, is an excellent conductor because the valence electron is very weakly bound to the individual atom. In the crystal structure of the solid metal these electrons are no longer associated with individual atoms, but wander freely through the crystal. This does not happen

in silver bromide because the valence electrons lost by silver atoms are strongly bound to bromine atoms.

In the theory of the formation of the latent image, the following steps are assumed to take place:

- A photon of light imparts its energy to an outer-shell electron of one of the bromine ions. The electron is freed from its attachment to that ion and becomes a conduction electron; i.e., it can move through the crystal structure.

- The resulting neutral bromine atom eventually makes its way to the surface of the crystal, where it is absorbed by the gelatin

- The conduction-band electron wanders through the crystal until it reaches a place where its progress is interrupted or slowed down. This may occur at an interior defect in the crystal, at an impurity atom, or, most significant for latent image formation, at the surface of the crystal.

- The electron then combines with a silver ion to produce an atom of metallic silver. Existing atoms of silver then serve as powerful attracting centers for subsequent electrons.

Experimental work has shown that a single atom of silver in a microcrystal of silver halide is ordinarily not sufficient to sensitize the grain (or make it <u>developable</u>). However, as few as <u>four</u> or <u>five</u> silver atoms will suffice. Such a grouping of silver atoms at the grain surface is called a <u>sensitivity speck</u>. The latent image is formed by the pattern of sensitized grains within the emulsion. In areas of shadow, few grains receive enough light to become sensitized, while in brightly illuminated parts of the image most of the grains acquire sensitivity specks. Since the presence of a

small number of silver atoms will result, during the development process, in the conversion of the entire grain to metallic silver, there is an amplification factor that can be as high as 10^9. This is perhaps the most significant characteristic of the photographic process, explaining the high sensitivity of silver halide materials.

Under ideal conditions, the latent image is quite stable. Images can be recovered even after a lapse of many years. But a number of factors can lead to loss or degradation of the latent image, either through:

- Fogging -- that is, sensitization of the entire emulsion
- Desensitization -- the ionization of metallic silver to silver ions with the resulting loss of active sensitivity specks.

Of special concern in aerospace applications is the fogging of photographic film by x-rays or nuclear radiation (particles and gamma rays).

Prior to the development process, there is no practical way to detect the presence of a latent image in a photographic emulsion by chemical or physical means, since only a few atoms per grain are involved. A latent image can be detected only by developing it.

Development and processing of photographic materials - Development is a chemical process in which

- Sensitized grains of silver halide are decomposed to form metallic silver
- Unsensitized grains are not affected.

Development is an example of an important class of chemical processes known as reduction. They are usually studied along with the reverse reaction, oxidation. The processes of oxidation and reduction are now reviewed briefly.

When a metal like silver gives up its status as an element to enter into a chemical compound, like silver bromide, the outer (valence) electron of the silver atom gives up energy as it becomes chemically bonded to another atom. The process involving this energy transfer is called oxidation. (Note that this process need not involve the element oxygen.) Similarly, the process of supplying energy to break the chemical bond (between silver and bromine, for example), and restore the valence electron to the silver atom, is called reduction.

Since the silver halides are relatively unstable compounds, it is not difficult to find reducing agents that will decompose them and release metallic silver. Such agents are not useful as photographic developers, of course, because they do not distinguish between sensitized and unsensitized grains. A photographic developer is a reducing agent that is

- Weak
- Selective

in that it can affect only those silver halide grains that have been sensitized. Examples of the most widely used developers are given in Table 3.2-4. All of these are chemically similar, being based on a single benzene-ring structure, derived from phenol (carbolic acid). The chemistry of the development process has been thoroughly studied and is well established. It involves the transfer of electrons from molecules of the reducing agent to silver ions in the exposed halide grain, starting at the sensitivity speck or specks on the

TABLE 3.2-4

TYPICAL PHOTOGRAPHIC DEVELOPERS

TRADE OR COMMON NAME	CHEMICAL NAME
HQ	Hydroquinone
Metol	N-methyl-p-aminophenol
Phenidone	1-phenol-3-pyrazolidone
Glycine	N-(4-hydroxy phenyl)

surface and working into the interior of the grain (Fig. 3.2-4), until the whole grain has been reduced to a clump of metallic silver. The constituents of a typical developer solution include:

- The developing agent itself (see Table 3.2-4)
- An alkali, or accelerator, such as sodium carbonate, borax, or sodium hydroxide
- A preservative, such as sodium sulphite
- A restrainer, such as potassium bromide.

Following the completion of the development process, a silver halide solvent such as sodium or ammonium thiosulphate is used to remove unexposed (and hence undeveloped) crystals of silver halide from the photographic emulsion. Although silver halides are not soluble in water, they are soluble in a solution of sodium thiosulphate, usually called hypo by photographers. Extensive washing follows the fixing, or fixation, process, to remove residual silver halide/thiosulphate complexes, as well as the thiosulphate itself. The washing step is necessary since such chemicals, if left in the emulsion,

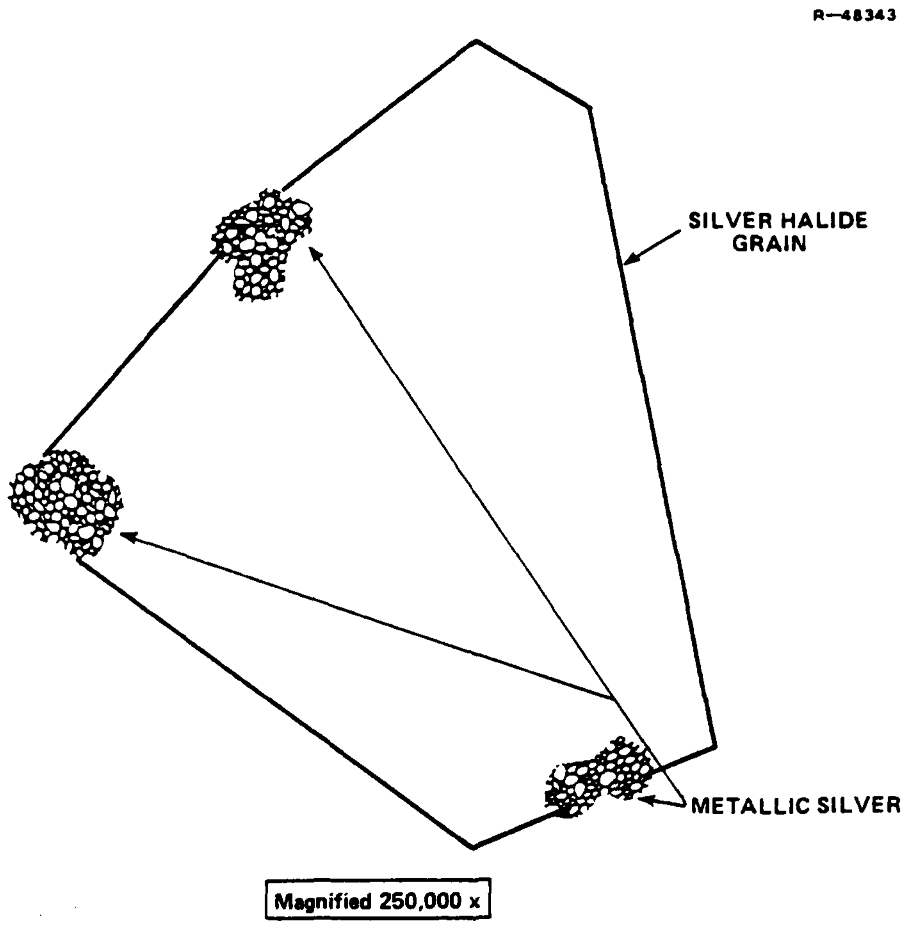


Figure 3.2-4 Developing Grain of Silver Halide

would eventually break down and destroy the photographic image by attacking the silver, a process analogous to tarnishing. The term archival processing refers to very thorough fixing and washing, often with the aid of additional chemical agents to enhance the solubility of the silver halide/thiosulphate complexes. Archivally processed images are very stable if protected from unfavorable environmental influences that would attack either the silver, the gelatin, or the base.

Properties of film and images - The response curve (or transfer function) characterizes a particular photographic material by expressing its output vs. input properties. The input is the exposure to light, measured in units of light intensity striking the material, multiplied by exposure time. The output is the degree of darkening produced in the photographic emulsion after development and processing. It is customary to express both input and output on logarithmic scales. For this reason, the degree of darkening is measured in terms of density, defined as the negative logarithm of the transmission ratio:

$$D = -\log_{10} T \tag{3.2-2}$$

Thus, for example, a film that transmits one-half of the incident light (T = 0.5) has a density of

$$D = 0.3 \tag{3.2-3}$$

while a film that transmits one hundredth of the incident light (T=0.01) has a density of

$$D = 2.0 \tag{3.2-4}$$

Input is measured in lux-seconds (or in foot-candle seconds in older publications). The response curve, called an H and D curve*, has the general form shown in Fig. 3.2-5, where the most important features are identified.

The film has a minimum density, known as base plus fog, because the base material absorbs some light (base density), and a certain small amount of development (fog) occurs

*After Hurter and Driffield, pioneer photographic scientists.

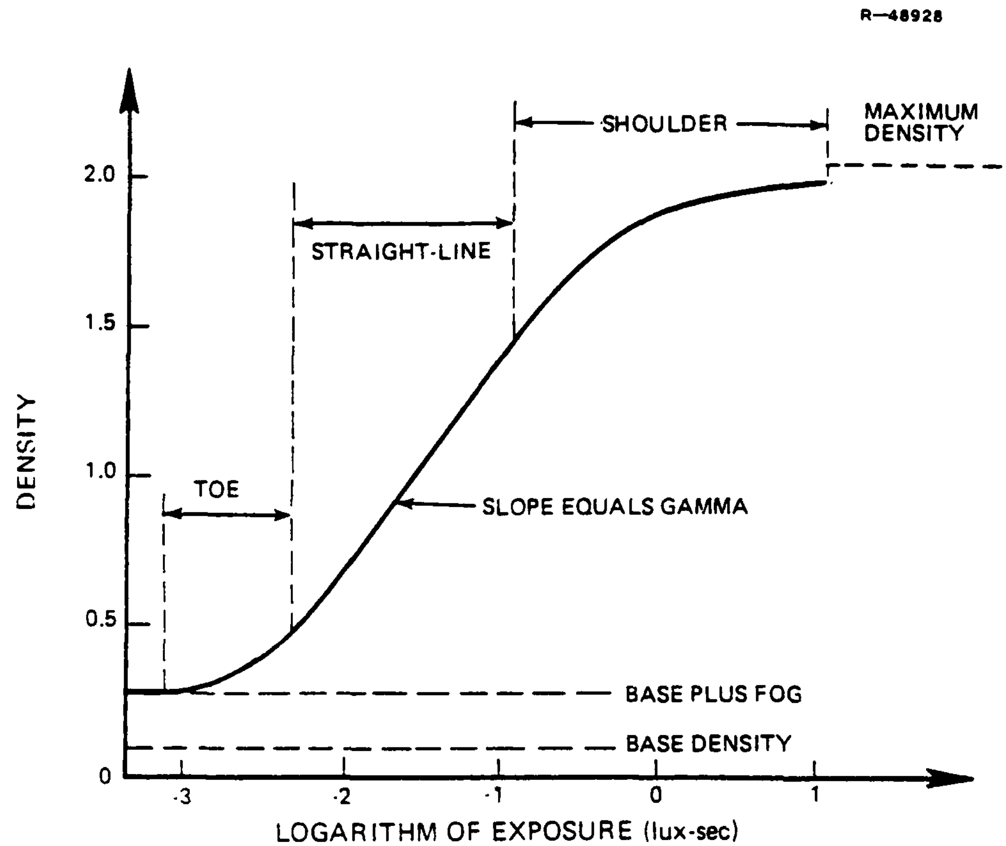


Figure 3.2-5 Example of the H and D Curve for Photographic Film

even in unexposed areas. The toe and shoulder regions represent exposure ranges in which the response is strongly nonlinear. For many applications this nonlinearity is useful; in pictorial and aerial photography, for example, the exposure time is adjusted to place the range of exposures on the toe and lower straight-line parts of the H and D curve (for optimal image quality). The slope of the straight-line part, called the gamma, is a measure of the contrast of the material. Although a gamma equal to one would appear to be ideal, reproducing the same range of light and dark in the image as appeared in the object, this is not true in practice. The brightness range in typical scenes is usually much too great to be matched by the available density range in the film;

the required compression is obtained by using materials with a gamma considerably less than one (0.5 to 0.7), as well as using the toe of the curve. Note that the form of the response curve is affected by the details of the development process (type of developer, dilution, time, temperature, etc.). Published H and D curves are based on standard processing as recommended by the film manufacturer; for critical applications, it is necessary to generate an empirical curve for the precise exposure and processing methods to be used.

Film speed is a single number expressing the relative sensitivity of different films by summarizing some of the important information contained in the H and D curve. Film speeds are used to set photoelectric exposure meters which estimate the optimal exposure time for a given scene. For black-and-white films used for general pictorial purposes, the film speed (called the ASA number or ASA speed) is defined by the equation

$$S = \frac{0.8}{H_m} \qquad (3.2\text{-}5)$$

where H_m (see Fig. 3.2-6) is the exposure required for a density of 0.1 greater than the base plus fog level, provided that the film is developed according to standard procedures. The standard development (Fig. 3.2-6) is based on the M and N points of the H and D curve, defined as follows:

- M -- the exposure H_m, as specified in the paragraph above
- N -- the exposure H_n, defined by the relation

$$\log H_n = \log H_m + 1.30 \qquad (3.2\text{-}6)$$

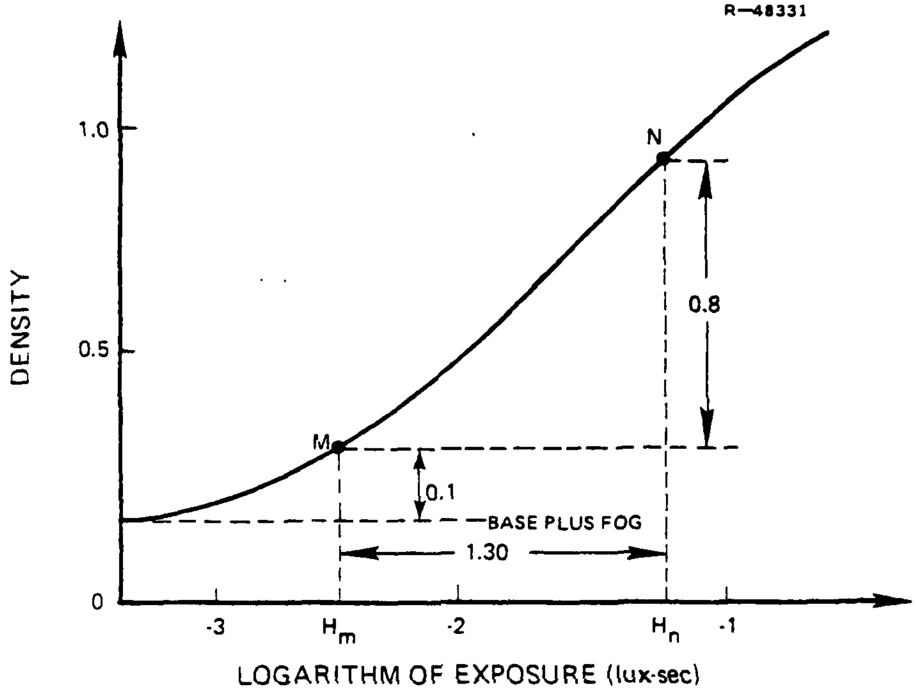


Figure 3.2-6   Determination of ASA Film Speed

Thus the N point corresponds to an exposure approximately 20 times greater than the M point. The development is standard if the density difference between these points is exactly 0.80. In practice, H_m corresponds to the minimum exposure giving just perceptible detail in dark shadows. For the case shown in Fig. 3.2-6 where

$$\log H_m = -2.5 \tag{3.2-7}$$

the ASA speed is

$$S = \frac{0.8}{H_m} = \frac{0.8}{0.00316} = 253 \tag{3.2-8}$$

This value corresponds to a film of moderate speed.

For aerial films, somewhat different speed numbers are used. Two systems are in effect, both illustrated in Fig. 3.2-7. The Federal Method B Speed, S_B, is defined as

$$S_B = \frac{0.5}{H_B} \tag{3.2-9}$$

where H_B is the exposure required for a density of 0.30 above the base plus fog level (Fig. 3.2-7). The Kodak Aerial Exposure Index, S_K, is given by the formula

$$S_K = \frac{0.5}{H_k} \tag{3.2-10}$$

where H_k is the exposure at that point on the toe of the H and D curve where the slope is 60 percent of gamma (Fig. 3.2-7). It is assumed in both cases that the film is developed according to the manufacturer's standard specification. For actual films, the two speed measures do not differ significantly.

Reciprocity refers to the assumption that the photographic effect of exposure to light is a function of the product of intensity and exposure time:

$$E = It \tag{3.2-11}$$

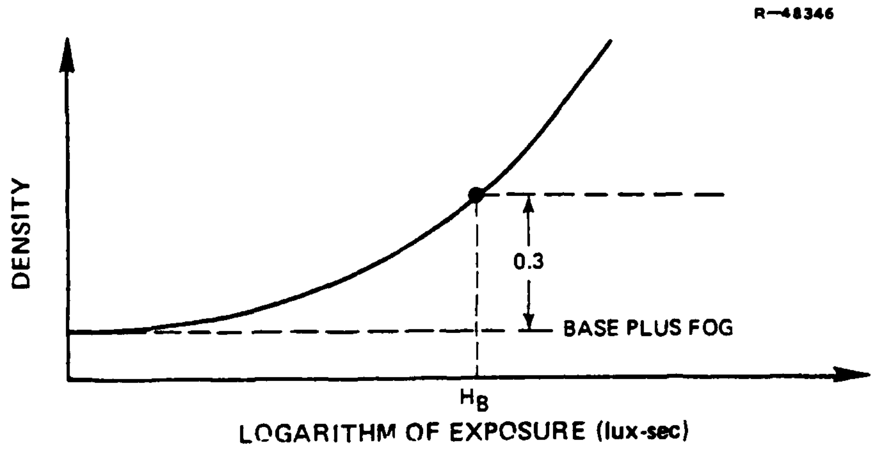
where

E is exposure in lux-seconds

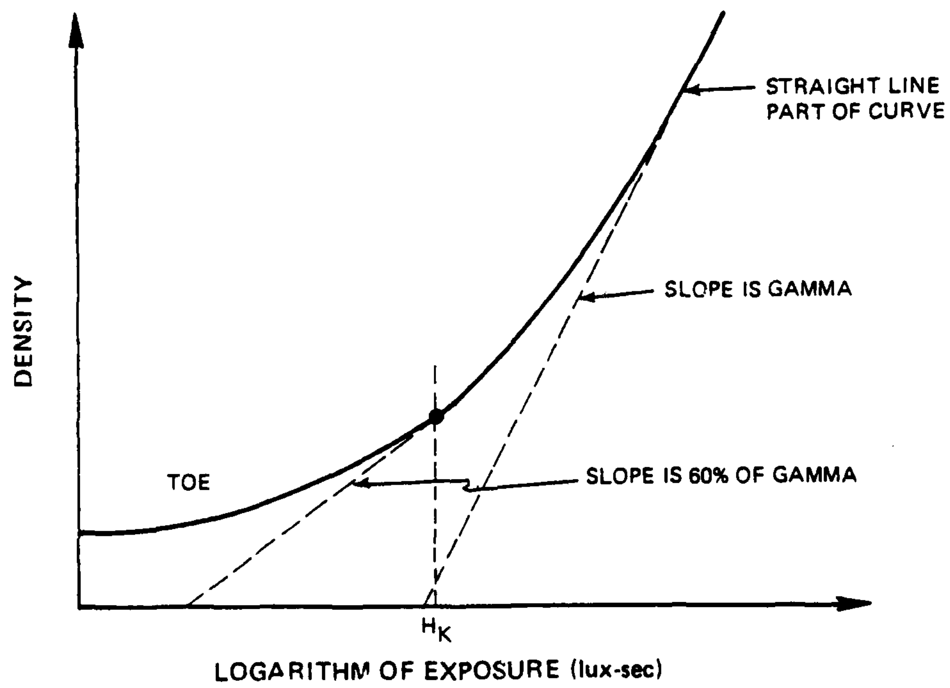
I is intensity in lux

t is exposure time in seconds

If the intensity is varying with time, the product in Eq. (3.2-11) must be replaced by an integral:



a) FEDERAL METHOD B SPEED FOR AERIAL FILM



b) KODAK AERIAL EXPOSURE INDEX

Figure 3.2-7 Determination of Aerial Film Speed

$$E = \int_{t_a}^{t_b} I(t)\, dt \qquad (3.2\text{-}12)$$

where

- $I(t)$ is the time-varying intensity
- t_a is the time when exposure begins
- t_b is the time when exposure ends

For constant illumination, the <u>reciprocity law</u> implies that all three of the following exposures produce the same photographic effect:

- 100 lux for 0.01 sec
- 10 lux for 0.1 sec
- 1 lux for 1 sec.

While the reciprocity law is a close approximation to the truth for normal light levels and exposure times, it fails to describe the actual behavior of photographic materials under conditions of:

- Extremely bright light for very short exposure times
- Extremely low levels of light for very long exposure times.

In both extreme cases, the resulting density is <u>less</u> than would be expected for the observed exposure in lux-seconds. Thus, to continue the example given above, where the total exposure was one lux-second, the following exposures might produce less density than would be predicted:

- 10^4 lux for 10^{-4} sec (high-speed electronic flash)
- 10^{-4} lux for 10^4 sec (photography by starlight).

This effect, known as reciprocity failure, is compensated for by exposure corrections based on data supplied by the film manufacturer.

Resolution, or resolving power, refers to the ability of the photographic material to reproduce fine detail. While the ultimate limit to film resolving power is set by the size and spacing of the developed silver grains that form the image, there are a number of other factors that combine to reduce resolution to a level far below the theoretical grain size limit. These include:

- Diffusion, or spreading, of the image by scattering within the emulsion
- Image spreading because of finite emulsion thickness
- Refraction of light within the emulsion
- Reflection of light from the rear surface of the film base (halation).

Resolving power is usually measured in units of line pairs per millimeter, and is based on the use of a test target with evenly spaced light and dark bands. A light band and its adjacent dark band form a line pair. In the practical determination of resolving power, a greatly reduced image of the test pattern is formed on the film. The developed film image is examined through a low-power microscope to determine whether the adjacent lines can be distinguished from one another. As

the image is made smaller and smaller, a minimum image size is reached at which the lines can no longer be resolved. The size of this minimum image determines the resolving power of the film.

It is clear that there is a subjective element in the measurement of resolving power, since the judgment of an observer is involved. In addition, there is a strong dependence on the contrast ratio between the light and dark lines. High-contrast patterns (where the light areas are 1000 times brighter than the dark areas) typically give resolving powers that are four to five times higher than those measured with patterns of very low contrast (lighting ratio of 1.25 to 1). Thus the resolving power of a film is usually stated for several different contrast ratios, and is based on the performance of an assumed average observer. As an example, Table 3.2-5 gives resolution data for a number of aerial films. Note the general rule that the faster films (those with higher exposure index or film speed) tend to have lower resolving power than slower films.

In high-quality aerial or satellite reconnaissance systems, it may be the resolving power of the film (rather than that of the lenses) that limits overall system performance. A knowledge of the relative target-image geometry then permits an estimate of the smallest features in the target that can be distinguished, based on the known resolving power of the film. The following example uses parameter values appropriate to satellite camera systems like those installed in Skylab.

Assume that the film width is 150 mm, with a resolution (for high-contrast features) of 100 line pairs per mm,

TABLE 3.2-5

RESOLUTION DATA FOR AERIAL FILMS

T-3840

FILM TYPE	AERIAL EXPOSURE INDEX	RESOLVING POWER (LINE PAIRS PER MM)	
		CONTRAST RATIO 1000:1	CONTRAST RATIO 1.6:1
Kodak Plus-X Aerographic Film 2402 (Estar Base)	80	100	50
Kodak Tri-X Aerographic Film 2403 (Estar Base)	250	80	20
Kodak Double-X Aerographic Film 2405 (Estar Base)	125	80	40
Kodak Plus-X Aerecon Film 8401	80	100	50
Kodak Panatomic-X Aerial Film 3400 (Estar Thin Base)	20	170	65
Kodak Plus-X Aerial Film 3401 (Estar Thin Base)	64	100	40
Kodak High-Definition Aerial 3414 (Estar Thin Base) 1414 (Estar Ultrathin Base)	2.5	630	250
Kodak Infrared Aerographic Film 2424 (Estar Base)	100	80	32
Kodal Aerochrome Infrared Film 2443 (Estar Base) 3443 (Estar Thin Base)	10	63	32
Kodak Ektachrome MS Aerographic Film 2448 (Estar Base)	6	80	40
Kodak Aerocolor Negative Film 2445 (Estar Base)	32	80	40
Kodak Aerial Color Film, (Estar Thin Base), SO-242	2	200	100
Kodak Ektachrome EF Aerographic Film (Estar Base) SO-397 (Estar Thin Base) SO-154	12	63	32
Du Pont Type 111R Aerial Negative (Cronar Base)	75	125	50

and that a strip of terrain 500 km wide is imaged in each exposure. Then, by considering similar triangles, it is seen that the dimensions of objects just barely resolvable (as determined by limitations of the photographic film) would be of the order of 35 m.

Another way to express resolving power of the entire photographic system is in terms of the limiting angular resolution. This is the angle subtended at the camera by the smallest distinguishable object. For the example just given, the angular resolution is about nine $\widehat{\text{sec}}$.

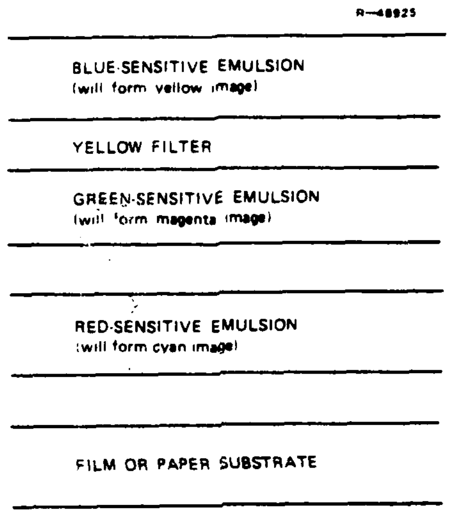
A more precise approach to image quality and resolution is discussed in a later section, under the heading The Modulation Transfer Function.

Materials and processes for color photography - Color photography is based on a property of the human eye: most visible colors can be reproduced satisfactorily by mixing the proper proportions of three basic, or primary, colors. Almost all color reproduction systems involve two steps:

- The formation of three separate images according to the distribution of light in the three main regions of the spectrum: blue, from 400 to 500 nm; green, 500 to 600 nm; and red, 600 to 700 nm
- The color synthesis stage, producing each of the three images in the appropriate color and combining them.

It is possible that the three images can be made on separate pieces of film, one after another, through appropriate filters, if the object is stationary and not changing. In graphic arts and printing applications, this is known as making color separation negatives. It is also possible to make the three negatives simultaneously, by using three optically aligned cameras or a single camera with a beam splitter and filter arrangement behind the lens. Systems like this have been used in satellite remote-sensing applications. Most of the time, however, the three images are made on a single piece of film with three photosensitive layers, selectively sensitized to the three regions of the spectrum, forming a simple reproduction system called the integral tri-pack. Almost all modern color photographic materials are based on this format.

The diagram in Fig. 3.2-8 and the following explanation refer to color negative material. The top layer is blue-



PRIMARY COLORS	COMPLEMENTARY COLORS*
BLUE (450 nm)	YELLOW
GREEN (550 nm)	MAGENTA
RED (650 nm)	CYAN (blue-green)

*Note: Each complementary color represents the result of removing the corresponding primary color from white light.

Figure 3.2-8 Structure of Color Negative Material

sensitive, and will form the yellow image. The middle layer is green-sensitive, and will form the magenta image. The bottom layer is red-sensitive, and will produce a cyan image. Underneath the blue-sensitive layer there is a yellow filter, which blocks blue light from reaching the two layers below, but allows green and red light to pass.

In order to obtain dye images from light-sensitive silver halides, a reaction known as chromogenic development is utilized. In this reaction, the exposed silver halide is developed in a special developer -- a widely used color developer is 4-aminoaniline or a derivative thereof. During the

reduction of the exposed silver halide image to metallic silver, the developing agent is oxidized. This oxidized developing agent reacts with a second component called a color former or color coupler, which may be incorporated in the developer solution or built into the emulsion, to form the actual dye. In color materials, the silver image is then removed by a process called bleaching, leaving only the dye image. The process just described must be followed by fixing and washing to remove unreacted silver halides and traces of processing chemicals. The result is a color negative, in which both colors and tonal values are reversed. A second exposure, onto a similar three-layer material coated on paper, produces a positive image or print.

Color reversal systems, which produce positive slides or transparencies in a single process (as distinguished from a negative-positive process), are now described. Exposure produces a latent negative image, which is developed in a black and white developer to form a negative silver image in each of the three layers. Color coupling takes no part in this reaction and no dye is formed. This stage leaves in the film, in addition to the negative silver image, a positive image of undeveloped silver halide, which can now be converted into a positive color image by chromogenic development. The key step, following initial development of the negative image, is reversal. The existing negative image must be removed, or destroyed, by means of chemicals that will dissolve the metallic silver but leave the unexposed and undeveloped silver halide intact. Once the negative image has been removed, the remaining silver halide is made developable either by a uniform exposure to light, or, more usually, by treatment with a chemical sensitizer, sometimes called a fogging agent. This positive image is now redeveloped by the use of a color-coupler developer system. A blue dye image is formed in the blue-sensitive

layer; a green image, in the green-sensitive layer; and a red image in the red-sensitive layer. The final step is the bleaching and removal of the positive silver image, leaving only the dye images in the film.

Color photography as just described is an example of multiband sensing, in which simultaneous images are recorded in various regions of the spectrum. Another important example is a three-layer material, commonly called color infrared film, in which two of the layers record zones of the visible spectrum, while the third layer is sensitized to the infrared. The resulting false color images (so-called because the infrared is developed as a visible color) are widely used in remote sensing and photo-interpretation. Even more powerful as a tool for interpretation are multilens cameras -- for example, the Itek Corporation produces a nine-lens aerial camera, recording in spectral bands ranging from ultraviolet, through visible, into the infrared. The nine images can be superimposed in various ways to produce color and false-color images, or examined individually for measurement of reflectivity and spectral characteristics.

Materials and processes for instant photography - There are various modifications of the standard photographic process designed to give rapid access to the developed images. One such system, often employed for document copying and photographic printing, uses sensitive materials having a developing agent, in an inactive form, incorporated in the emulsion. Development begins when the material is immersed in a strongly alkaline solution to activate the developer, which completes its action within a few seconds. Instead of fixation and washing, an alternative process called stabilization is usually used for rapid-access materials. This is done to save time and eliminate the need for running water. The price paid is

that the image is less stable; it will eventually fade or be discolored. Stabilization involves treating the undeveloped silver halide with appropriate chemicals to form silver halide complexes that are:

- Insensitive to light
- Transparent
- Stable enough to give the material a reasonable lifetime.

The entire process may be completed in ten seconds or less.

Another approach to rapid processing is known as <u>diffusion transfer</u>. This exists in a number of forms, two of which are used as examples:

- The Polaroid process
- The Kodak Bimat Transfer System.

In the Polaroid process, the negative is developed by a viscous mixture of developer and halide solvent. A sheet of receiving material is pressed into contact with the negative during development. As the silver image develops within the negative, the left-over undeveloped silver halide (constituting a positive image) diffuses into the receiving layer, where it is reduced to metallic silver by chemical agents called <u>development nuclei</u>. When processing is complete (10 to 15 seconds for some materials), the positive image is stripped away and is ready for use. In the original Polaroid process the negative was formed on paper and was not designed for further use; there is now, however, one type of Polaroid material in which the negative, on a transparent base, can be retained and used later for printing or enlargement.

The Kodak Bimat process has been adapted for in-flight processing of long rolls of aerial film. The two components are:

- The negative film
- The Bimat transfer film, a receiving material (provided with development nuclei) coated on a transparent base.

The Bimat material is presoaked in processing solution, which has no effect until the two films are pressed into contact during development (Fig. 3.2-9). The diffusion transfer process then forms a negative in the original film, and a positive image in the Bimat transfer layer. Both films are spooled

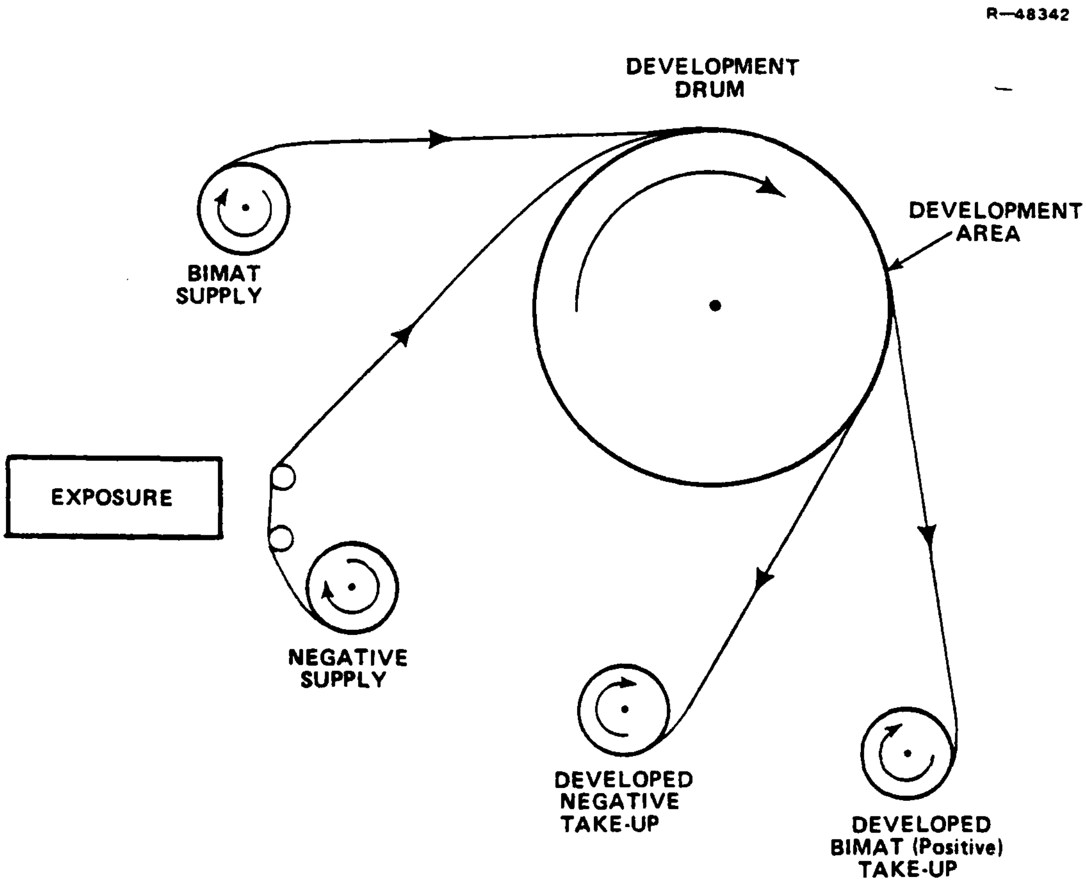


Figure 3.2-9 Schematic of the Bimat Development Process

onto take-up reels after leaving the development drum. They are ready for immediate use, but require washing (to remove traces of processing chemicals) if long-term retention is required. The Bimat positive is convenient for light-table examination by photo-interpreters; the negative, on the other hand, is usually filed for archival purposes, and used to make paper prints and enlargements.

Besides the convenience of immediate access to the image, diffusion transfer systems have the additional advantage of using less silver, thus making more efficient use of this expensive and potentially scarce resource. They are, of course, also usable under conditions where water is not available for conventional processing.

Dye diffusion systems, involving the formation of dye images in three separate layers, and the migration of these images onto a receiving sheet, form the basis for instant color photography (Polaroid, Kodak). A detailed treatment of instant color photography is beyond the scope of this text.

3.2.2 Electro-Optical Imaging Systems

Electro-optical imaging systems convert the radiant energy forming an image into patterns of electrical charge that can be stored, manipulated, and converted into visible form. The optical systems that form the image are similar to those in photographic cameras; the essential difference is that the image is formed, not on film, but on a photosensitive surface, called a photocathode, in which photons of incident light transfer their energy to electrons of the cathode's material. Electro-optical devices have become important for these reasons:

- Instant availability of the image, either for direct viewing or in electrical form for storage, analysis, or transmission

- Greater sensitivity to low levels of light, along with the possibility of electronic image intensification

- Linearity of the input-output relation over a wide range of light intensity (unlike photographic film, which is strongly nonlinear), facilitating specialized image analysis and interpretation

- Possibility of very wide spectral response.

The principal disadvantage of electro-optical devices, when compared to photographic imaging, is <u>limited resolution</u>.

<u>Properties of photocathodes</u> - Photosensitive surfaces may be characterized according to

- Spectral response -- there are three main categories (Table 3.2-6)

- Physical mechanism by which they operate -- the principal types are

  - photoconductive surfaces

  - photoemissive surfaces

<u>Photoconductive devices</u> are based on the physical process of photoconductivity, in which the electrical resistance of a suitable semiconductor material is <u>decreased</u> by illumination. Electrons in such materials are normally <u>bound</u> and unable to move through the crystal structure. Absorption of a photon of light energy provides sufficient energy to free an electron and permit it to move easily through the crystal, acting as a charge carrier (<u>conduction</u> electron). With additional mobile

TABLE 3.2-6

SPECTRAL PROPERTIES OF PHOTOCONDUCTIVE SURFACES

CATEGORY	PROPERTIES
Far Ultraviolet	Peak response at about 130 nm, with long wavelength cutoff at 200 nm -- no response to visible light
Solar-blind	Sensitive to ultraviolet, with long wavelength cutoff at 400 nm -- relatively insensitive to solar radiation transmitted through the atmosphere
Visible/Infrared	Ultraviolet, visible light, and infrared up to 1000 nm

charge carriers available to respond to an applied voltage, the effective resistance of the material decreases. Hence the consequences of increased illumination on the photoconductive surface are:

- More photons absorbed by bound electrons
- More mobile electrons to serve as charge carriers
- More charge transfer (current) in response to an applied voltage
- Lower electrical resistance.

Among the devices based on photoconduction is the vidicon, widely used for television purposes.

Photoemissive devices are based on the phenomenon of the photoelectric effect, in which electrons are ejected from the surface of the photosensitive material. These photoelectrons can be accelerated and focused by electrostatic or electromagnetic fields, and used to create visible images or electrical

signals in a variety of ways. Examples of devices based on photoemission are the image orthicon and image isocon, frequently used in television.

Imaging Device Principles - Most electro-optical imaging devices involve two fundamental stages:

- Image formation
- Image readout by raster scan, generally involving a focused electron beam.

The concept of a raster scan is shown in Fig. 3.2-10. United States commercial television standards, for example, call for scanning the image as 525 discrete lines. For reduction of flicker, the scanning is interlaced -- first the odd-numbered lines are scanned (in 1/60 second), and then the even-numbered lines are scanned. The complete image is sampled, therefore, 30 times a second. Regardless of the quality of the lens system, the associated electronics, and other components of the system, the ultimate limit to image resolution is set by the number of scanning lines. For special applications, devices with far higher scan rates and many more lines per image have been developed.

In order to move an electron beam across the photosensitive surface on which an image has been formed, it is necessary to:

- Focus electrons into a narrow beam
- Deflect the beam to trace out the raster scan pattern.

Each of these requirements can be met by the use of electrostatic fields, formed by suitably shaped charged electrodes,

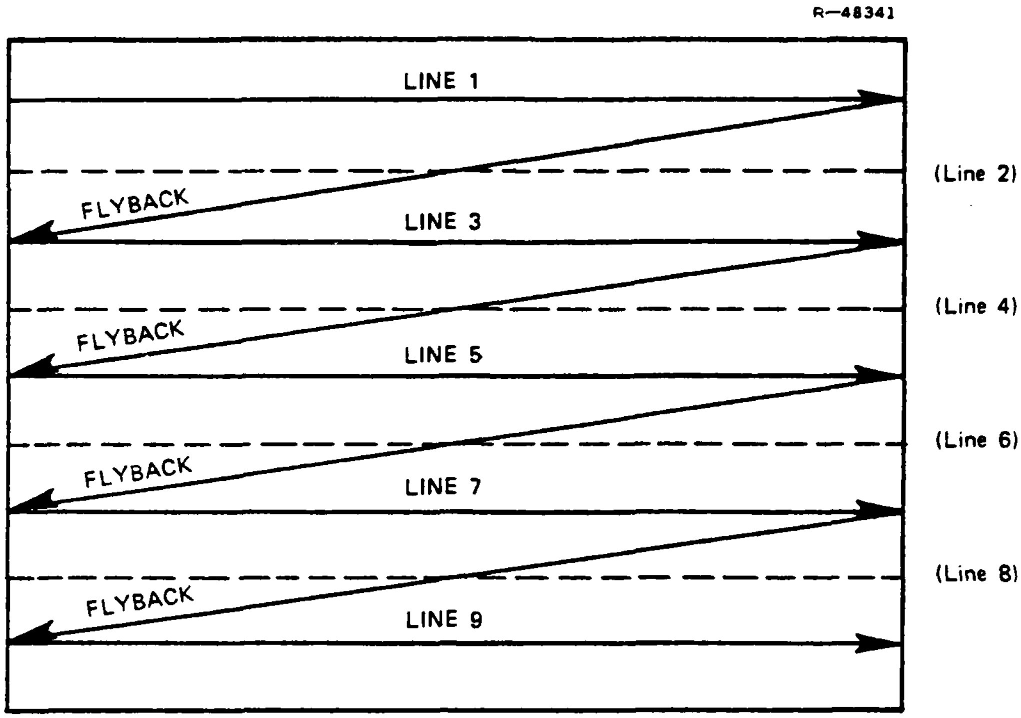


Figure 3.2-10 Concept of the Raster Scan

or by electromagnetic fields, formed by properly configured and spaced electromagnetic coils.

Television Camera Tubes - As important examples of electro-optical imaging devices, two widely used types of television camera tubes will be described briefly. These are the vidicon, an example of a photoconductive device, and the image orthicon, an example of a photoemissive device.

The vidicon has the general configuration shown in Fig. 3.2-11. Facing the lens system, there is a transparent conducting film on the surface of the vidicon's photocathode, connected to the signal electrode (output terminal). Beneath this layer is the photoconductive material itself, generally several microns in thickness. The basic principle of operation

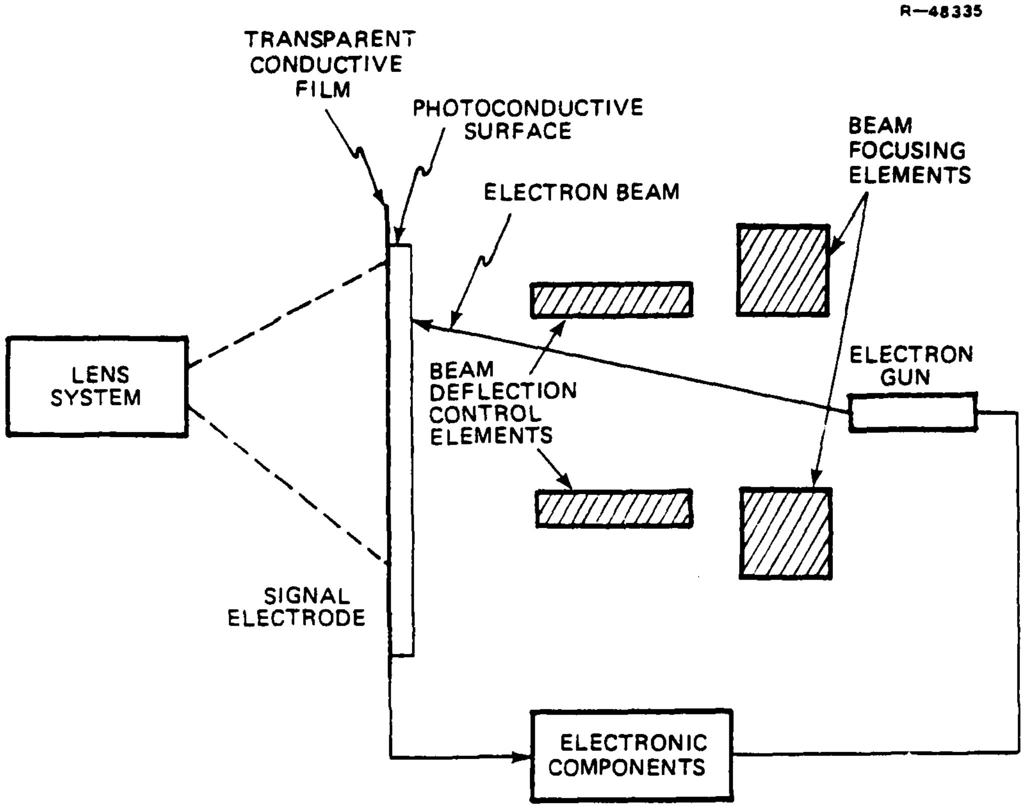


Figure 3.2-11 Configuration of the Vidicon

is that the photoconductive layer acts like a leaky capacitor.* This capacitor is uniformly charged by the passage of the scanning electron beam. It is then partially discharged, to a different extent at each point of the image, in accordance with the light-induced conductivity at that point.

As the electron beam scans over a particular point, it restores the charge lost by leakage. An electric current, equal to the rate at which charge is being restored, flows through the signal electrode and the associated electronic circuitry. Thus the input sensed by the electronic components, at any instant of time, depends on the charge being restored

*A capacitor with some resistance in parallel with it.

at the image point where the scanning beam is then located. This depends, in turn, on the charge lost at that point by leakage, which is proportional to the light intensity there.

Figure 3.2-12 illustrates this concept further by considering two points -- one brightly illuminated, the other totally dark. Both points start with a uniform potential difference across the capacitor formed by the photoconductive surface and the signal electrode (Fig. 3.2-12a). At the brightly illuminated point, conduction electrons are available, leading to a relatively low electrical resistance (Fig. 3.2-12b), while there are relatively few conduction electrons in the dark area. As a result (Fig. 3.2-12c), charge leaks away at the illuminated point, reducing the potential there. At the dark point, the resistance of the photoconductive surface remains high. Very little of the charge leaks away, and the potential remains close to the original uniform scanning potential.

As the electron beam scans past the first point (3.2-12d), it must restore all of the lost potential. While it is doing this, there is a substantial current in the circuit consisting of the electron beam, the photoconductor capacitor, the signal electrode, and the remaining electronic circuitry. When the beam scans the dark point (Fig. 3.2-12e), there is very little charge to be restored, and, consequently, very little current in the circuit at that time. Thus the image information is transferred into the form of vidicon signal current intensity as a function of time.

<u>The Image Orthicon</u> consists of these main components:

- The image section
- The storage target

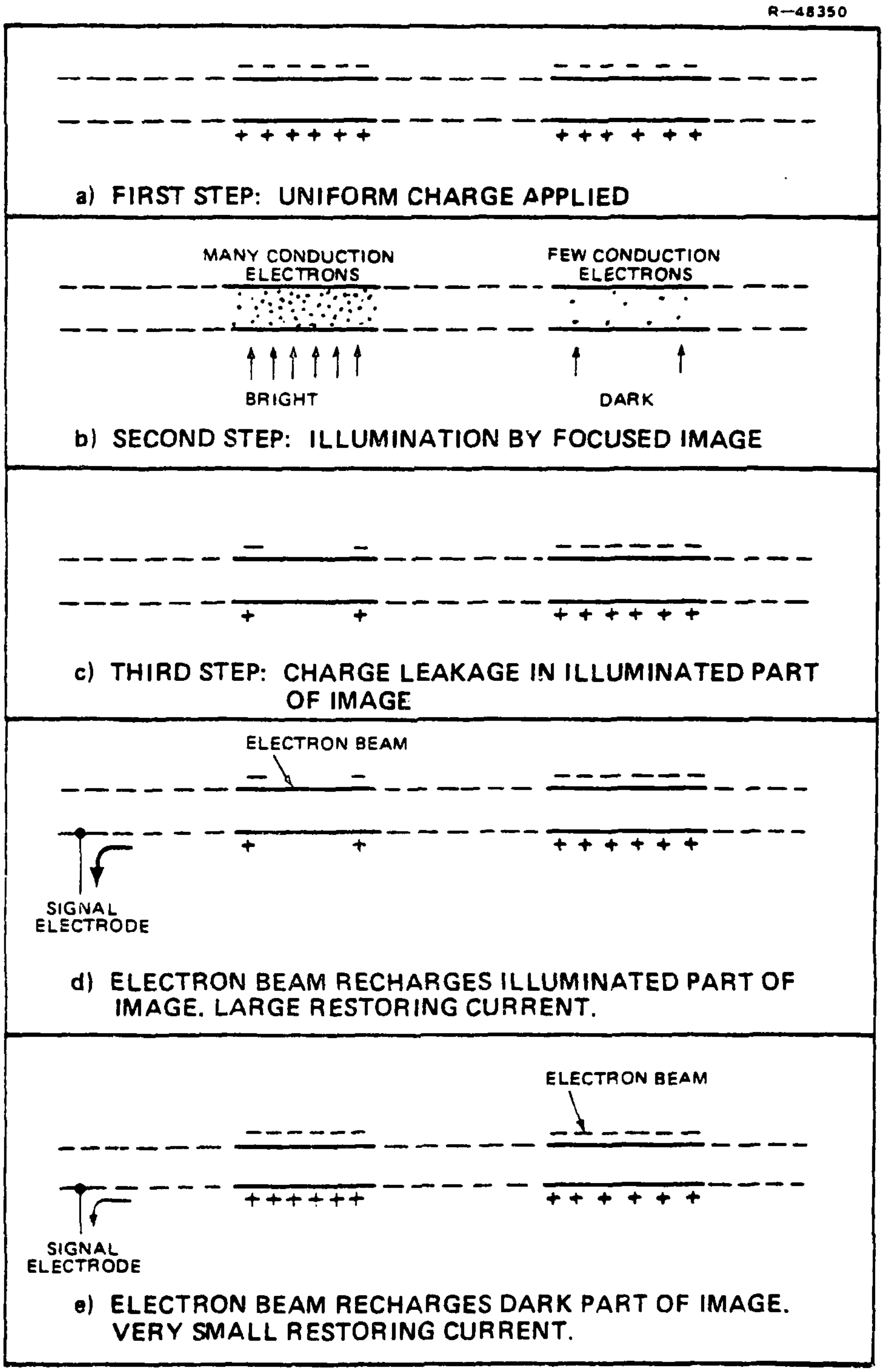


Figure 3.2-12 Operation of the Vidicon

- The scanning electron beam
- The electron multiplier, which amplifies the return beam signal.

These components are illustrated in Fig. 3.2-13.

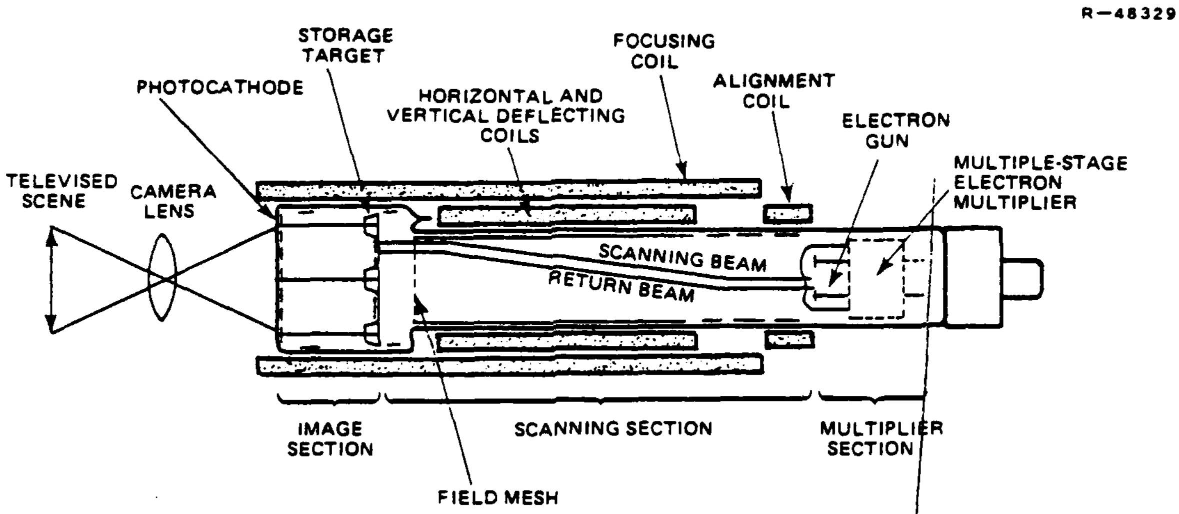


Figure 3.2-13   Components of the Image Orthicon

An image is formed by the lens system on the photocathode. Electrons released in brightly illuminated areas by the photoelectric effect are accelerated by uniform electrostatic forces generated within the image section, to strike the storage target at high velocity. At this point an amplification process occurs, through the phenomenon of secondary emission. For each electron that strikes the storage target, several electrons are ejected from the target surface. The multiplication factor (number of electrons emitted for each incident electron) depends on the details of device fabrication, accelerating voltages used, and target material. A typical

value of the multiplication factor is five. Thus, as shown in Fig. 3.2-14, for each electron that strikes the storage target, five are ejected, with a net effect of the loss (depletion) of four electrons, resulting in a positive charge of four electron units at that point. There has been an amplification by a factor of four, and an image in the form of positive charges is left on the surface of the storage target.

The readout mechanism is a low-velocity scanning beam, focused and deflected, in a raster scan pattern, by electromagnetic fields in the scanning section (Fig. 3.2-13). As the beam scans a point in the storage target, it deposits enough electrons to neutralize the positive charge at that point. Remaining beam electrons are turned back and refocused to form the return beam, which is directed into the multiplier section (Fig. 3.2-13). The return beam is amplitude modulated in accordance with the charge pattern on the storage target, and hence represents the variations of light intensity in the image.

The electron multiplier section uses secondary emission to amplify the return beam which bears the image signal. The amplified beam is then applied to the electronic circuitry associated with the image orthicon. The two stages of amplification, in the image section and in the multiplier section, are responsible for the high sensitivity of the image orthicon.

Color television, which has important technical and military applications (in addition to its entertainment function), is generally implemented by forming three images. The separate images are obtained by using beam splitters and appropriate color filters, in the primary colors* red, green,

*The choice of primary colors (a set of colors from which the others can be generated) is not unique and is dictated by the spectral characteristics of available filters. The reader may also be familiar with red, yellow, and blue as a primary color set.

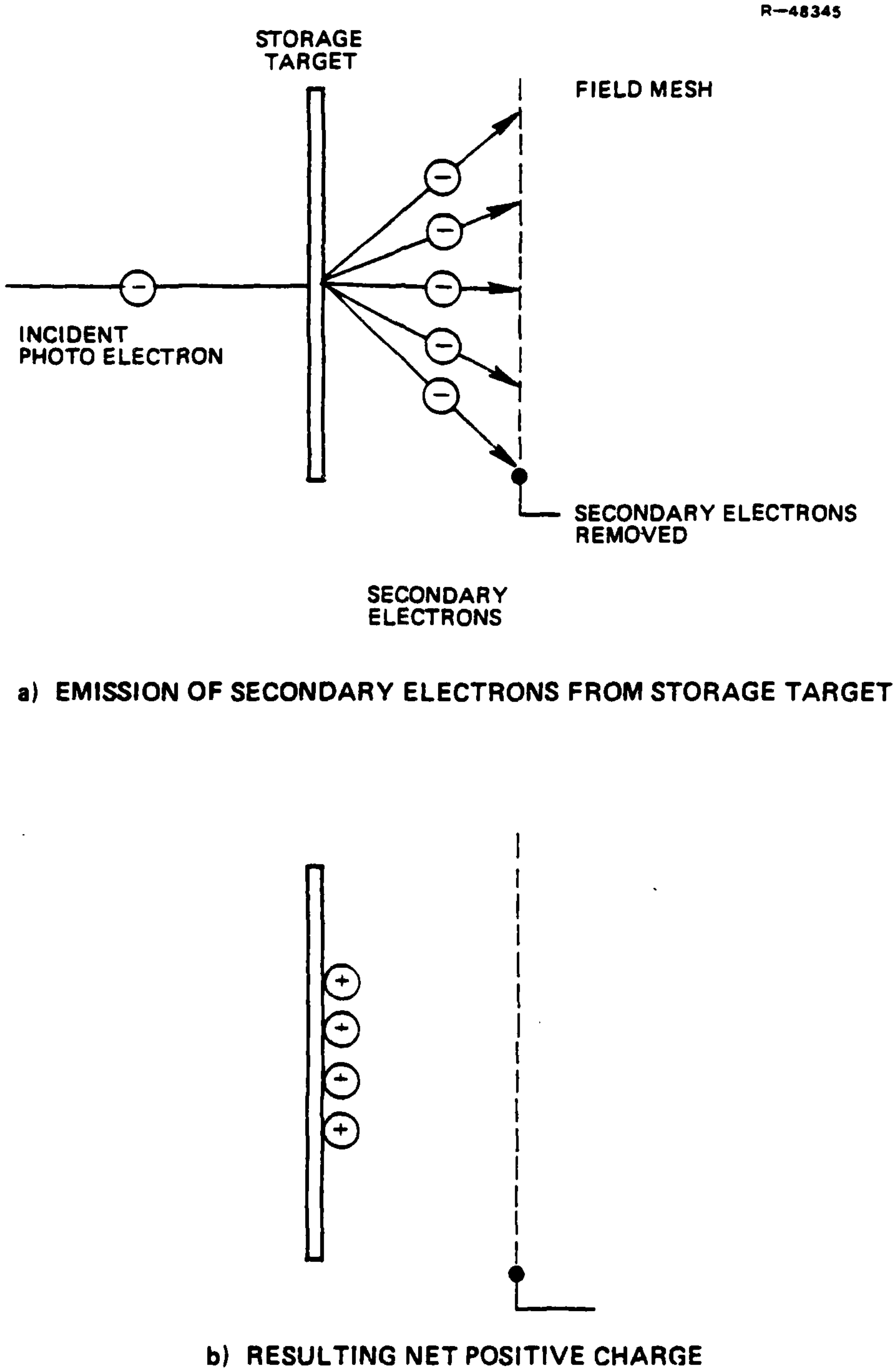


Figure 3.2-14   Secondary Emission at the Storage Target

and blue. Each image is scanned with a conventional camera tube. Usually a variation on the vidicon principle called the plumbicon is used for this purpose. The combination of lens system, beam splitter, filters, three plumbicons, and associated electronic circuitry can be made into a remarkably compact package.

Performance of television tubes is usually characterized in terms of

- Transfer function - expressing the electrical output as a function of the light input (Fig. 3.2-15) analogous to the H and D curve for photographic film
- Spectral response - describing the relative sensitivity of the device to various wavelengths of light (Fig. 3.2-16)
- Resolution - given in various ways:
  - total number of scan lines
  - line pairs per millimeter

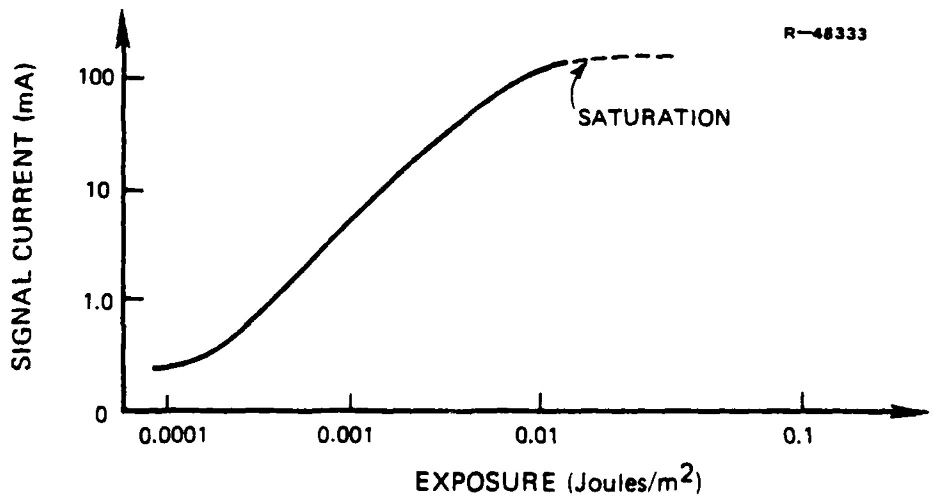


Figure 3.2-15 Transfer Function for a Typical Television Camera Tube (Vidicon)

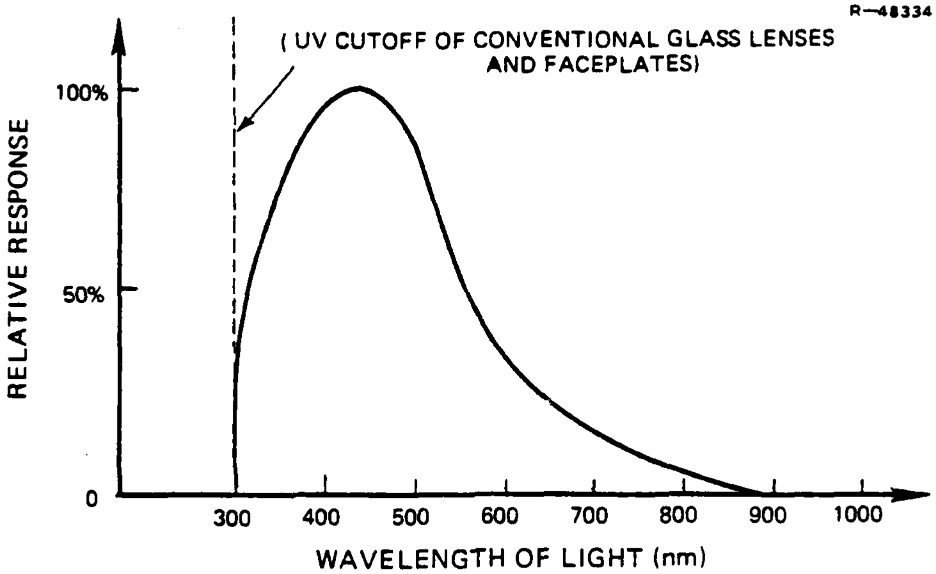


Figure 3.2-16 Spectral Response for a Typical Television Camera Tube

- total number of pixels (picture elements) resolved
- modulation transfer function (MTF), to be discussed in a later section.

Other electro-optical devices - Image intensifiers are widely used for viewing under dim-light conditions. There are numerous military, scientific, and law-enforcement applications. The operation of an image intensifier can be described in terms of the following functions:

- Conversion of the image (UV, visible light, or infrared) into an electron emission pattern
- Intensification of the electron image by an amplification process
- Conversion of the intensified electron image into a visible image.

For example, the original image may be formed on a photoemissive surface. The resulting photoelectrons are focused and accelerated by electrostatic or electromagnetic fields, and then bombard the rear surface of a luminescent phosphor, which glows in proportion to the electron flux incident upon it. The visible image may be

- Viewed directly, or through a lens system
- Photographed
- Picked up by a conventional television camera tube
- Used as the input to a second intensifier (cascaded intensifiers).

Storage Tubes are used to retain images for long periods of time, until they are required for readout and analysis. There are many important applications in reconnaisance and information-handling systems. Tubes of vidicon design can be adapted for use as storage devices in the following way:

- Initially scan the photoconductive surface to apply a uniform distribution of electrical charge
- Expose the image for a limited period of time. This results in partial discharge depending on the intensity of illumination at each point of the image. The photoconductive surface is then shielded from light during the storage period. Charge neutralization depletes the population of photoconduction electrons within the sensitive surface, leading to a restoration of the original high-resistance condition, and the pattern of charge distribution is reasonably stable for storage times of up to several hours.
- Reactivate the raster scan mechanism to read out the stored image and prepare the photoconductive surface for its next exposure.

Infrared imaging devices exist in a variety of forms. Two important examples will be mentioned. The thermicon is an infrared version of the vidicon, in which an image is formed by an infrared optical system (using mirrors, or lenses fabricated from infrared transmitting materials like calcium fluoride) on a heat-sensitive variety of photocathode. Other infrared imaging devices are of the direct-view type, analogous to the image intensifer, with the visible image formed by the bombardment of a phosphor screen by photoelectrons originating at the photoemissive surface. Military, law-enforcement, intelligence, engineering, and scientific applications of devices that "see in the dark" are manifold.

3.2.3 Digital Image Characteristics

General Description - In nature, visual scenes are usually continuous distributions of energy in three spatial coordinates at a given time. However, when these scenes are imaged by an optical system onto a sensor, the result is a two-dimensional image which is often formed in discrete steps of the continuous variables. This is done for one of two reasons: either the imaging device is actually made up of an array of small sensors (e.g., the retina of the human eye or a photodiode array), or the processing of the imagery is done by a digital computer which can operate only on discrete numbers.

It has been found experimentally that the human visual system has a spatial frequency cutoff at about 60 cycles/degree. Thus, any details with higher spatial frequencies are not discernable to a human viewer. In addition, the hardware mechanisms for producing and displaying digital images have a strong impact on their structure. Section 3.3.3 contains a detailed description of the sampling process and the hardware required for it. Essentially, hardware exists for creating digital images from

their continuous counterparts in sizes ranging from a few picture elements (pixels)* to hundreds of millions of pixels, with the number of quantization levels ranging from two to a thousand. Hard-copy or photographic representations of the digital data can be produced from most scanning devices by replacing their light detectors with light sources. The light sources generate an image which is developed to provide the copy.

It has been found that a level of detail of at least 512 by 512 pixels per image is typically required for a standard 4 by 5 inch digital image to appear to the human visual system as nearly equivalent to its continuous counterpart at normal viewing distances. Larger images require a proportionally higher number of pixels. It should be noted that even a 512 by 512 pixel image with 256 quantization levels requires a digital storage capacity of over 2 million bits. Larger images increase storage requirements by the square of the number of pixels per line.

Because the hard-copy displays require some type of photographic processing, the results cannot be seen in real time. This has led to the development of CRT (Cathode Ray Tube) soft-copy displays based on standard television technology. Here, the digital image is displayed for the viewer on the face of the CRT in real time. However, current technology supports image sizes ranging from about 256 by 256 pixels to 1024 by 1024 pixels, with 512 by 512 being the most prevalent. Viewing distances must be approximately five to six times the picture height to eliminate the eye's sensitivity to rastering (the discrete nature of the scan lines). This makes the raster frequency greater than the eye's spatial frequency cutoff.

---

*A more precise definition is given in Section 3.3.2.

The result of these tradeoffs is that the usual digital image is 512 by 512 pixels with 256 quantization levels (requiring 8 bits to express). Larger images are often divided into contiguous blocks of this size for display and processing. This is usually a good compromise, since the viewing distances can be adjusted to provide an almost continuous appearance found satisfactory by most observers.

Image Statistics - It is an unfortunate fact that the probability distributions of normal imagery vary considerably from scene to scene. No general representation exists. Histograms representing the probability densities of typical imagery often fall into one of two categories: either the number of dark pixels greatly exceeds the number of middle gray and bright values (see Figure 3.2-17), or there are peaks in the dark and bright areas as compared to middle gray. Although these variations cannot be characterized or parameterized by any known probability distributions, they at least imply that the choice of simple source models (e.g., Gaussian) is usually not valid.

Better models for image description are currently the subject of considerable attention in the image processing community, because they are the basis for theoretical analysis of the properties of images.

3.2.4 Non-Silver Imaging Materials

Except for electro-optical systems (e.g., television), most materials for recording images are based on the extreme light-sensitivity of the halide compounds of silver. Many other substances are sensitive to light, at least to some extent, and there has always been considerable interest in the possibility of developing practical photographic systems that do not use silver. Non-silver photographic systems are of very great potential importance for two main reasons:

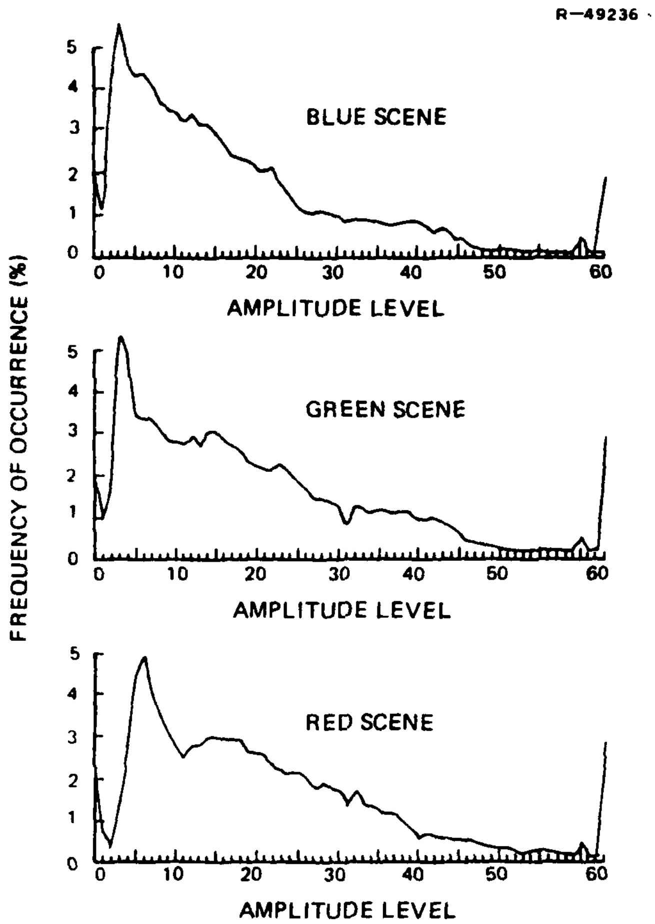


Figure 3.2-17 Typical Histograms of Red, Green, and Blue Tristimulus Values of a Color Image

- Silver is expensive, and the world's available supply of this important industrial metal may eventually be exhausted

- Non-silver materials may have special characteristics that are useful in particular applications.

Table 3.2-7 lists the most common categories of non-silver imaging materials, some of which are useful at the present

TABLE 3.2-7

CLASSIFICATION OF NON-SILVER IMAGING MATERIALS

Non-silver inorganic chemicals:
• Iron compounds
• Copper compounds
• Dichromated colloids
• Lead halides
• Thallium halides
• Chalcogenides
• Mercury compounds
Thermographic systems
Photochromic systems
Diazo systems
Polymeric systems
Electrophotographic (or electrostatic) systems

time in specialized application areas like document copying and the graphic arts. Others are still in the experimental stages of development.

Two of these systems -- diazo and electrostatic -- are reserved for detailed discussion in the next section. The remainder of the non-silver systems are now briefly reviewed.

Non-silver inorganic systems - Historically, some of the oldest of the non-silver systems make use of the light-sensitivity of compounds of iron. The familiar blueprint process, once widely used for the reproduction of engineering drawings, is an example. Two limitations to the wider application of systems based on iron are

- Low sensitivity, requiring long exposures to very bright light
- High contrast, limiting the ability of such systems to reproduce the intermediate shades of continuous-tone images.

<u>Copper</u> belongs to the same group of elements as silver in the periodic table. It might, therefore, be expected that copper compounds -- for example, the copper halides -- should show a photosensitivity analogous to that of silver compounds. Since copper is less expensive than silver and more widely available, there have been many attempts to develop copper-based photographic systems. So far, no commercially usable system has emerged. One problem is the difficulty in finding a suitable fixing agent. Another is that the photographic speed is very low, requiring long exposures under extremely bright light.

<u>Dichromated colloids</u> are a category of photosensitive materials currently used in a variety of photographic and photomechanical (printing) processes. Colloids are materials of organic origin like gums (for example, gum arabic), albumin (egg white), and gelatin. When they are treated with solutions of potassium dichromate (also called bichromate), the resulting complexes are sensitive to light. A latent image is formed by bright light (preferably UV) to produce areas of varying solubility, since the effect of the light on the dichromate-colloid complex is to harden it (a process analogous to the tanning of leather). One way to make the latent image visible and permanent is by washing in warm water, which dissolves away the unhardened colloid and leaves the hardened areas intact. After drying, these hardened areas may be treated with a dye, which is absorbed, thus creating an image in color.

Another commercially important application of dichromated colloids (dichromated gelatin, for example) is the preparation of photoresists. The hardened gelatin that remains after the washing process is used as a protective layer in a subsequent etching operation. Photoresists are used in the graphic arts for making printing forms of every description, such as letterpress, halftone blocks, planographic printing plates, gravure plates and cylinders, and in the silk screen process. Silk screening, usually thought of as a tool for print-making by creative artists, is used to manufacture a wide variety of items ranging from street signs to electronic printed circuits. The areas of hardened gelatin (exposed to light) block the flow of ink through the silk-screen stencil, while the unhardened gelatin (in areas receiving no light) is washed away, permitting ink to flow through these areas.

Tests of the halide compounds of lead reveal a limited sensitivity to light. Attempts to base photographic systems on the use of lead halides are still in the experimental stages. Thallium halides are also photosensitive. One major disadvantage blocking their possible practical use is the extreme toxicity of the materials involved.

Chalcogenides (sometimes called chalcogenide glasses) are glass-like materials formed by controlled melting, under vacuum conditions, of cadmium sulphide mixed with small quantities of sulphides of arsenic, selenium, and lead. The resulting materials are sensitive to light in limited portions of the spectrum. Chalcogenide systems are still experimental.

Mercury compounds have interesting photographic properties. The halides, in fact, strongly resemble silver halides in many ways. Mercury compounds have long been used for certain photographic applications (the original Daguerreotype, for

example) but their use is very much restricted because there is no known practical way of fixing the image to make it permanent. Toxicity is, of course, also a problem with materials containing mercury. Sensitivity to light is lower than in the case of silver, and is limited to the UV and blue end of the spectrum.

Thermographic systems - Thermography refers to any process of imaging in which heat is used to bring about a visual effect. Practical applications have generally been limited to document copying -- for example, the Thermofax process. In some applications, a heat-sensitive paper is placed in contact with the printed material to be copied. Intense illumination (heat and infrared) causes heat to be transmitted to the copy paper in the printed (dark) areas of the original, resulting in the darkening of the copy in these areas. In earlier forms of thermographic copy systems, the sensitive papers consisted of a dark-colored base coated with a white fusible material. Heat would melt away the white coating to form a dark image. Modern thermographic papers generally involve two or more chemical components that are inert at room temperature but react, when heated, to form an intensely colored compound.

Photochromic Systems - Photochromism refers to a reversible color change when many inorganic or organic compounds are exposed to light. Photochromic glass has, for example, been used for so-called automatic sunglasses. Unlike silver halide films, photochromic materials are completely grain free, and permit very high resolution levels -- about 1000 lines per mm. They are, therefore, very well suited to the production of microfiche. Images can be reduced by factors as high as 150 to 1, and very high density information storage is thus achieved. Although many of the useful photochromic glasses contain a small amount of silver halide, they are classified as non-silver systems because the image is not formed by metallic silver.

Polymeric Systems - Polymeric systems are based on the phenomenon of photopolymerization, in which light (visible or UV) induces the relatively small molecules of the sensitive substance (called the monomer) to link together to form the giant molecules of the resulting polymer. A latent image is then formed by the regions of unpolymerized (original monomer) and polymerized material. Various differences in physical properties between the monomer and the polymer can be utilized to make the image visible or to transfer it to a permanent medium.

The most obvious difference is that monomers are typically liquids, while the corresponding polymers are either solids or extremely viscous liquids. This difference in state is generally accompanied by a difference in solubility. Hence unpolymerized areas of the image can be brushed away, or washed off with suitable solvents, leaving the polymerized areas intact. Photopolymers based on this principle are widely used in the preparation of printing plates, and as photoresists for the etching of printed circuits and integrated circuits.

Other physical properties upon which polymeric imaging systems are based include

- Adhesion
- Tackiness
- Viscosity
- Conductivity
- Refractive index.

These properties of the latent image are used in a variety of ways to fix or transfer the image. Most actual applications are in the graphic arts and printing trades.

At the present time, the use of polymeric imaging systems is limited by

- Extremely high contrast (unsuitable for reproduction of continuous tone images, except by means of halftone processes)
- Low light-sensitivity, requiring long exposure at high levels of illumination.

3.2.5 Electrostatic and Vesicular (KC) Film Systems

Two particular non-silver imaging systems with important applications are discussed in this section. These are electrostatic systems, as exemplified by the Xerox document copying process, and vesicular film systems -- for example, the Kalvar Corporation (KC) film process.

Electrostatic Systems - In an electrostatic (or electrophotographic) imaging system, a latent image is formed by a pattern of electrically charged and uncharged areas on a surface of photosensitive material. The image is made visible by use of a pigmented powder, bearing an opposite electrostatic charge. The basic sequence of steps is

- Uniform charge applied to photosensitive surface
- Exposure to light, with charge being dissipated in bright areas of image, retained in dark parts
- Development, by application of pigmented powder that adheres to charged areas
- Transfer of image to receiving material
- Fixing of image on receiving material
- Cleaning of remaining pigment from photosensitive surface.

The steps are now discussed in detail, with reference to Fig. 3.2-18.

Step a (See Fig. 3.2-18a) is known as charge spraying. A corona discharge is used to spray ions onto the surface of the photosensivity material. Fine wires held at a short fixed distance from the photoconductor are moved across the surface at a speed of about 5 cm/sec. The wires are maintained at a potential above 7000 volts. The high electric field at the wire surface ionizes air molecules near the wire. The ions move to the photoconductor surface and charge it, usually to several hundred volts.

In step b (Fig. 3.2-18b), the original is illuminated, and a lens system forms an image on the photosensitive surface. During this exposure, the electric charge is dissipated in those areas of the image that are brightly illuminated, and is not dissipated in the dark areas of the image. The mechanism is as follows: when a photon of sufficient energy is absorbed by the photoconductive layer, an electron is excited and moves freely (conduction electron). This electron neutralizes some of the previously applied uniform positive charge. Note that the photoconducting layer on which the surface charge was formed is not laterally conducting -- that is, adjacent charged and uncharged areas will not neutralize one another. The latent charge image can be retained for a considerable period of time.

Step c (Fig. 3.2-18c) converts the pattern of electric charges forming the latent image into a visible image. This may be done in a number of ways. Typical is the dry powder development method. Charged grains consisting of a pigmented thermoplastic are placed in contact with or brought near to the surface bearing the electrostatic image. These particles (referred to as the toner) are then attracted by

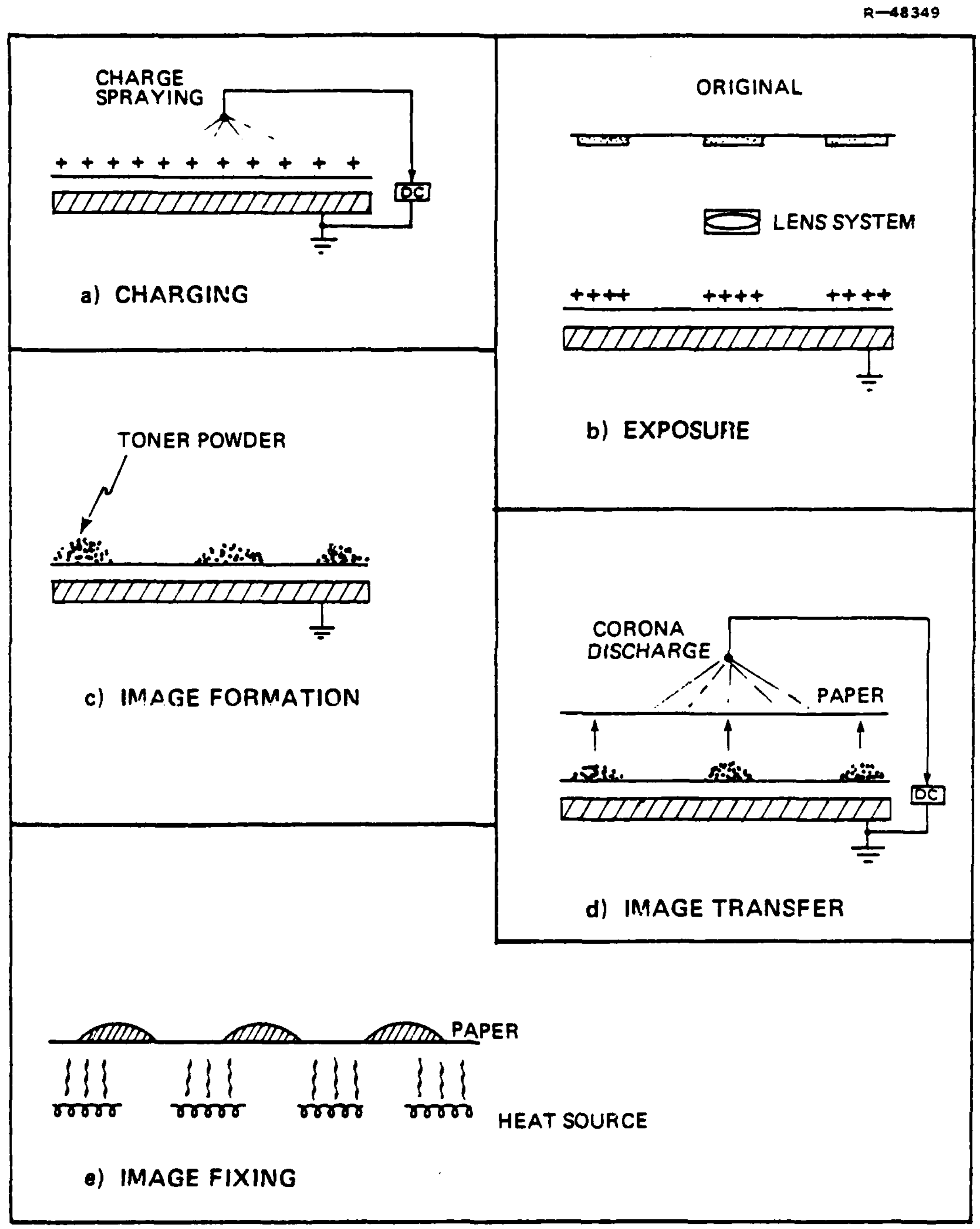


Figure 3.2-18 Steps in the Electrostatic Transfer Process

electrostatic forces to the charged areas of the image. Excess powder is removed by blowing or brushing off. A visible image is now formed on the photoconducting surface.

Step d (Fig. 3.2-18d) involves the transfer of the image to a receiving surface. In document copying applications this may be a sheet of plain paper. In printing applications, the image is transferred to the surface of a lithographic printing plate, which is subsequently inked and used to produce a number of copies. The receiving material is placed over the image, and a high electrical potential is applied, using the same corona discharge system that sensitized the receptor surface in the first place. Electrostatic attraction then moves the particles of toner from the photoreceptor surface to the receiving surface of the paper or printing plate. When the receiver is peeled away from the photoreceptor, it carries with it a large portion of the toner image.

In step e (Fig. 3.2-18e) the image is fixed (made permanent). This is done by the controlled application of heat to the receiver surface, causing the particles of toner material to melt together and become permanently attached to the receiver. In the final step (not illustrated), residual powder remaining on the photosensitive surface is removed by brushing or blowing. The surface is now ready for another cycle of image formation and transfer.

The design of the photosensitive surface is highly critical to the success of the electrostatic imaging system. A typical photosensitive surface, which may be fabricated either in planar or in cylindrical form, is made up of four separate layers:

- A top surface, the charged barrier layer, on which the image is formed
- A bulk charge transport layer -- e.g., selenium, a semiconductor
- A bottom barrier layer -- e.g., aluminum oxide (an insulator)
- A conductive substrate -- e.g., aluminum metal.

In addition to their widespread use in document copying and in the preparation of lithographic printing plates, electrostatic imaging systems are being applied in the preparation of hard copy from microfilm, the production of microimages, in photoengraving, radiographic recording (electroradiography or xeroradiography), printing from photographic negatives, computer output printing, rapid access display, and holographic recording. Numerous variations of the basic electrostatic imaging concept are in experimental or developmental stages, and may eventually be of commercial and military importance. It should also be noted that full-color electrostatic imaging systems are in production, based on multiple exposure through appropriate filters, and the use of colored toner powders.

Vesicular film systems - In vesicular processes, use is made of the light-scattering properties of small bubbles of gas (mostly nitrogen) formed in the interior of a clear thermoplastic layer by the photo-induced decomposition of a class of organic compounds called diazonium salts (diazo, for short). In its original form, the vesicular process* is negative working like ordinary photographic film.

---

*Often called the Kalvar or KC process (named for the Kalvar Corporation, which manufactures and markets vesicular film and related materials and equipment).

There are three basic steps in the formation of a vesicular image, as illustrated in Fig. 3.2-19:

- Exposure - UV radiation (most effective wavelength is 385 nm) induces the decomposition of diazo molecules, forming nucleating centers of nitrogen and other gases (Fig. 3.2-19a)
- Thermal development - application of heat causes nucleating centers to grow from small bubbles 0.5 to 5.0 micrometers in diameter into vesicles. Temperatures of from 80 to 150 C are applied for times ranging from milliseconds to five seconds or more (Fig. 3.2-19b).
- Fixation - to make the image permanent, it is necessary to eliminate the light-sensitive material that was not exposed and converted in the image-forming process. This is done by an overall exposure to UV radiation, followed by controlled gentle heating for extended periods of time, to allow nitrogen and other volatile products to diffuse out of the film without forming vesicles (Fig. 3.2-19c).

Once formed and fixed, the negative vesicular image is used like an ordinary photographic negative (Fig. 3.2-20). Positive vesicular images are used as transparencies for projection and viewing. An outstanding advantage of vesicular materials is their extremely high resolution -- of the order of 500 line pairs per mm -- making these materials suitable for microfilm and microfiche preparation, and for use as motion picture print film.

There are many other photographic applications of diazo materials. In most processes, the exposure to light and subsequent chemical development result in the formation of intensely colored dyes. Both positive materials (dye formed in unexposed areas) and negative materials (dye formed in

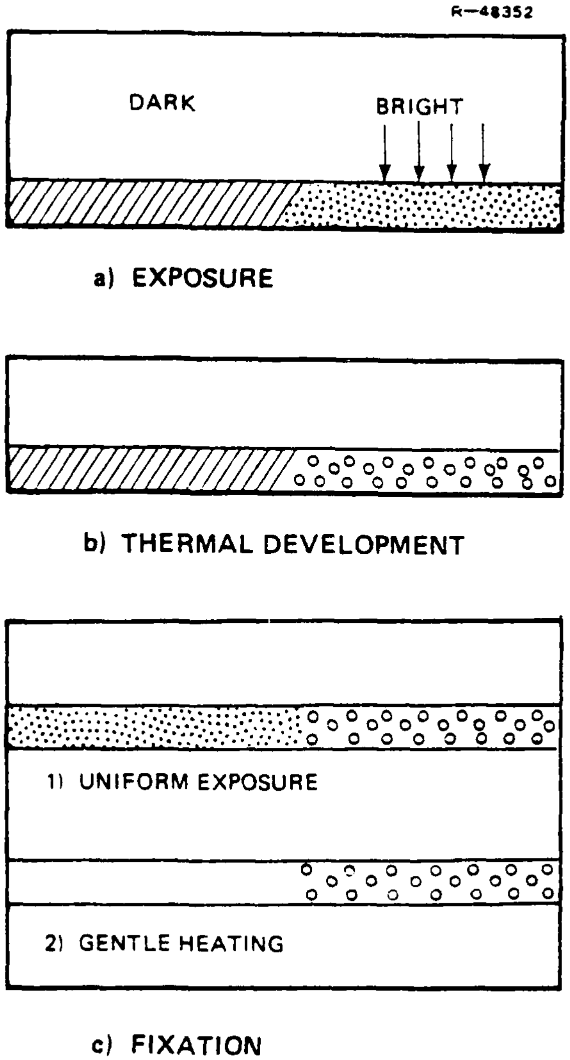


Figure 3.2-19 Formation of Vesicular Image

exposed areas) are in use. The dye images have practically no granularity; thus the resolution capacities of diazo materials are very high. A typical developed grain of silver halide may have a diameter of 300 nm, compared with 1.5 nm for a molecule of converted diazo-dye.

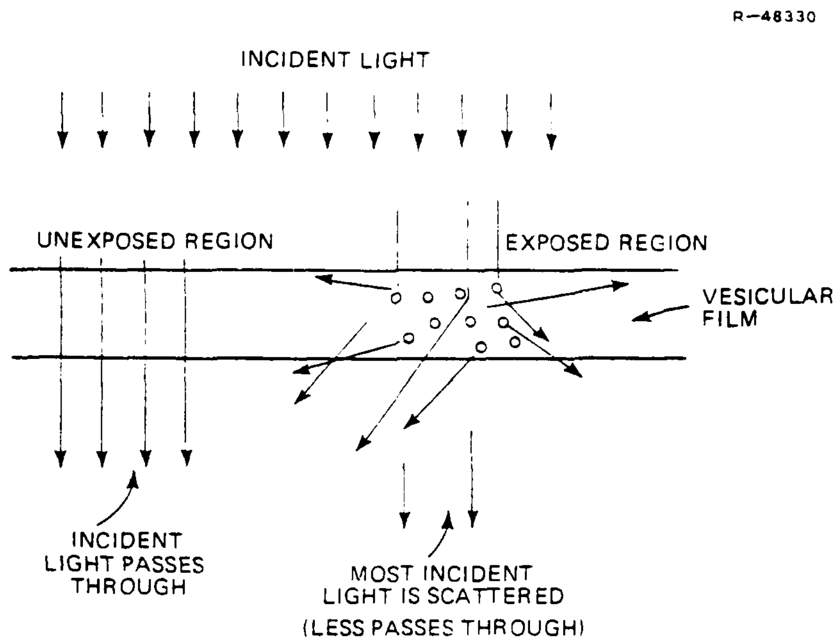


Figure 3.2-20 Use of Vesicular Image

3.2.6 Modulation Transfer Function

Several ways of characterizing the degree of resolution of an imaging system have been used in the preceding sections of this chapter. For example, two such measures are

- The angle subtended by the smallest object that can just barely be distinguished in the resulting image (typically measured in minutes of arc)
- The dimensions of the smallest pattern of regular light and dark lines that can be made out (lines, or line-pairs, per millimeter).

While such simple measures are useful, they fail to reflect the fact that resolution of fine detail tends to degrade continuously as the characteristic dimension decreases. There is really no precise point at which image response changes from

perfect to zero. Another disadvantage of simple resolution measures is that there is no way to combine the stated or measured resolutions of various components of an imaging system -- for example, camera lens system, film, enlarging lens system, print paper -- to predict the resolution of the overall system.

A more sophisticated approach to characterizing the quality of imaging systems is based on the same techniques of spectral (frequency) analysis used by electrical engineers to analyze electronic communication systems. Expressing the performance of an imaging system in terms of the <u>modulation transfer function</u> (MTF) is analogous to describing an audio amplifier or speaker system in terms of its frequency response curve.

Just as the frequency response analysis of an electronic component uses a sinusoidal test signal (single frequency) as input, the analysis of an optical system uses as the test input a pattern whose light intensity varies sinusoidally. Figure 3.2-21 represents the intensity variation along the cross-section of such a bar pattern, as described by the equation

$$T(x) = A + B \cos 2\pi f x \qquad (3.2\text{-}13)$$

where

- x = position in the pattern
- T = intensity of light transmitted or emitted
- f = spatial frequency of the pattern (usually measured in cycles per millimeter)
- A, B are constants characterizing the pattern

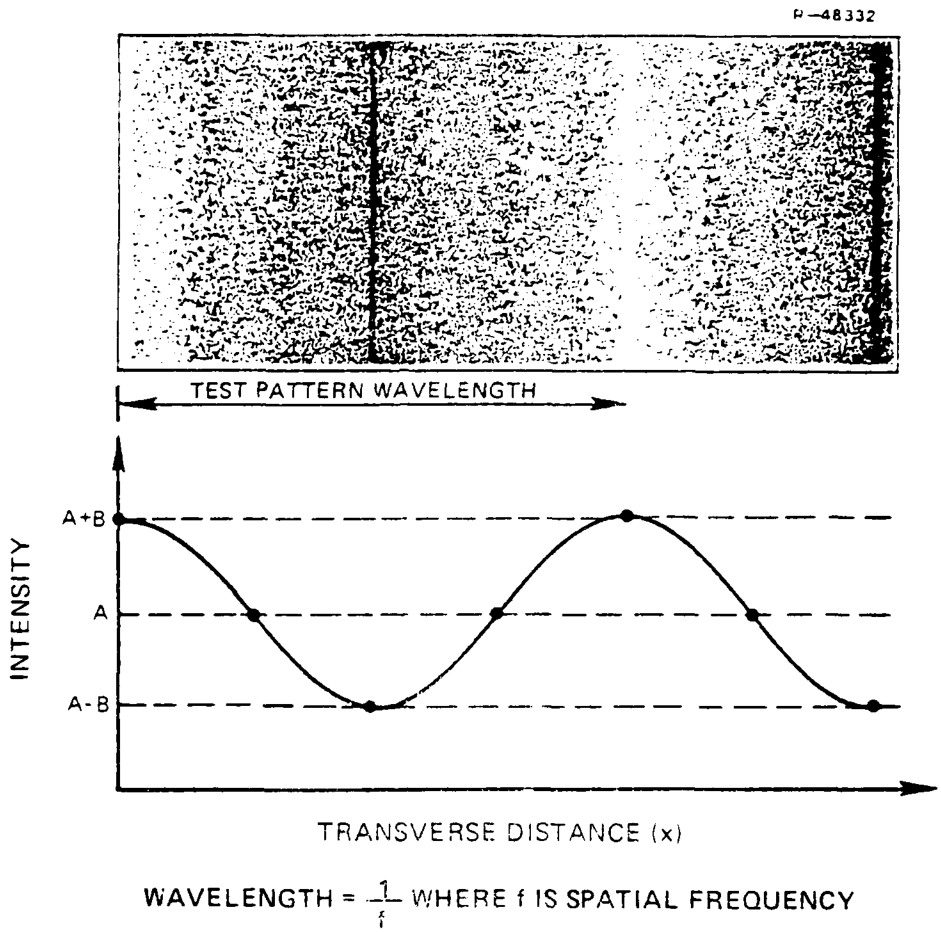


Figure 3.2-21 Standard Bar Pattern for Determining Modulation Transfer Factor

The ratio

$$M_o = \frac{B}{A} \qquad (3.2\text{-}14)$$

is called the <u>modulation</u> of the test object. The image is also a sinusoidally varying bar pattern, with light variations characterized by the equation

$$T_I(x) = a + b \cos 2\pi fx \qquad (3.2\text{-}15)$$

and an image modulation given by

$$M_I = \frac{b}{a} \tag{3.2-16}$$

Because the imperfections of the imaging system cause some of the light that should have been focused on the bright bars to spill over into the adjacent dark areas, the image modulation will always be less than the object modulation. The ratio

$$m = \frac{M_I}{M_O} \tag{3.2-17}$$

is called the modulation transfer factor at the spatial frequency f, characterizing the degree of degradation suffered by a bar pattern of that frequency. A curve or equation showing how the modulation transfer factor changes as a function of the spatial frequency is called a modulation transfer function, usually abbreviated as MTF.

Examples of MTFs are shown in Fig. 3.2-22. Figure 3.2-22a is the MTF of the human eye as an overall imaging system, including the properties of the lens and of the retina. Figure 3.2-22b is the MTF of a standard black and white aerial film; by itself, this MTF would represent the performance of an aerial camera with a theoretically perfect lens system, limited only by the resolution capabilities of the film. If combined with an MTF characterizing the lens system, this would then describe the performance of an actual camera system.

MTFs are combined by multiplying them together. Thus, for example, a reconnaissance system consisting of

- Telescopic optical system
- Infrared-sensitive image converting electro-optical system

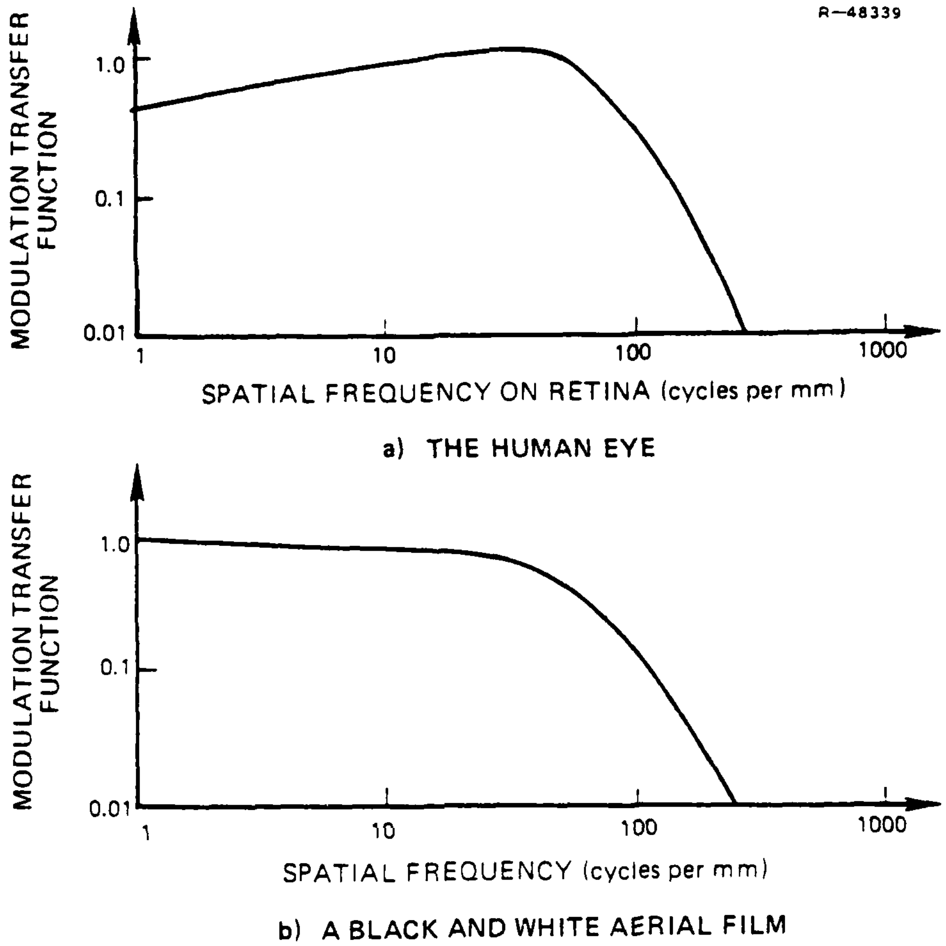


Figure 3.2-22 Examples of the Modulation Transfer Function

- Recording optical system
- Photographic film

is described by the product of four individual MTFs, each representing the performance of one component of the system.

(†)For the benefit of those readers with the requisite mathematical background, a more formal description of the

---

(†)The remainder of this section contains material at a more advanced level than the rest of the text.

MTF and its properties is now outlined. The treatment which follows is introductory. The reader who wishes to examine these issues more closely is referred to any of several texts on signal processing given in the reading list at the end of this Unit.

It is appropriate to begin by introducing the concept of a <u>point spread function</u> (PSF). This represents the response of the imaging system to a single point of light (a star, for example, or a very small pinhole). In a lens or system of lenses, the various imperfections (called <u>aberrations</u>) cause the image of a point source to be spread out into a (more or less) blurred spot of finite extent (Fig. 3.2-23a). Even if a lens had no aberrations, the wave nature of light would cause image spreading through the process of <u>diffraction</u>, resulting in a bright central spot surrounded by a series of fainter rings (Fig. 3.2-23b). If a perfect point of light were to be projected on photographic film, the result would be a blurred image because of scattering of light in the emulsion, reflection from the film backing, and the finite size of the silver halide grains (Fig. 3.2-23c). In an image tube (like a vidicon), spreading of the image results from the granular nature of the photosensitive surface, from charge leakage, and from the finite size of the scanning electron beam.

The PSF of an optical system provides sufficient information to predict the response of the system to <u>any</u> object, since the object can be regarded as a superposition of an infinite number of point sources. A one-dimensional object (i.e., a line source along which the intensity varies) is assumed in this discussion, in order to keep the mathematics as simple as possible. For the same reason it is assumed that the image is the same size as the object. Then the distribution of light in the image can be predicted from

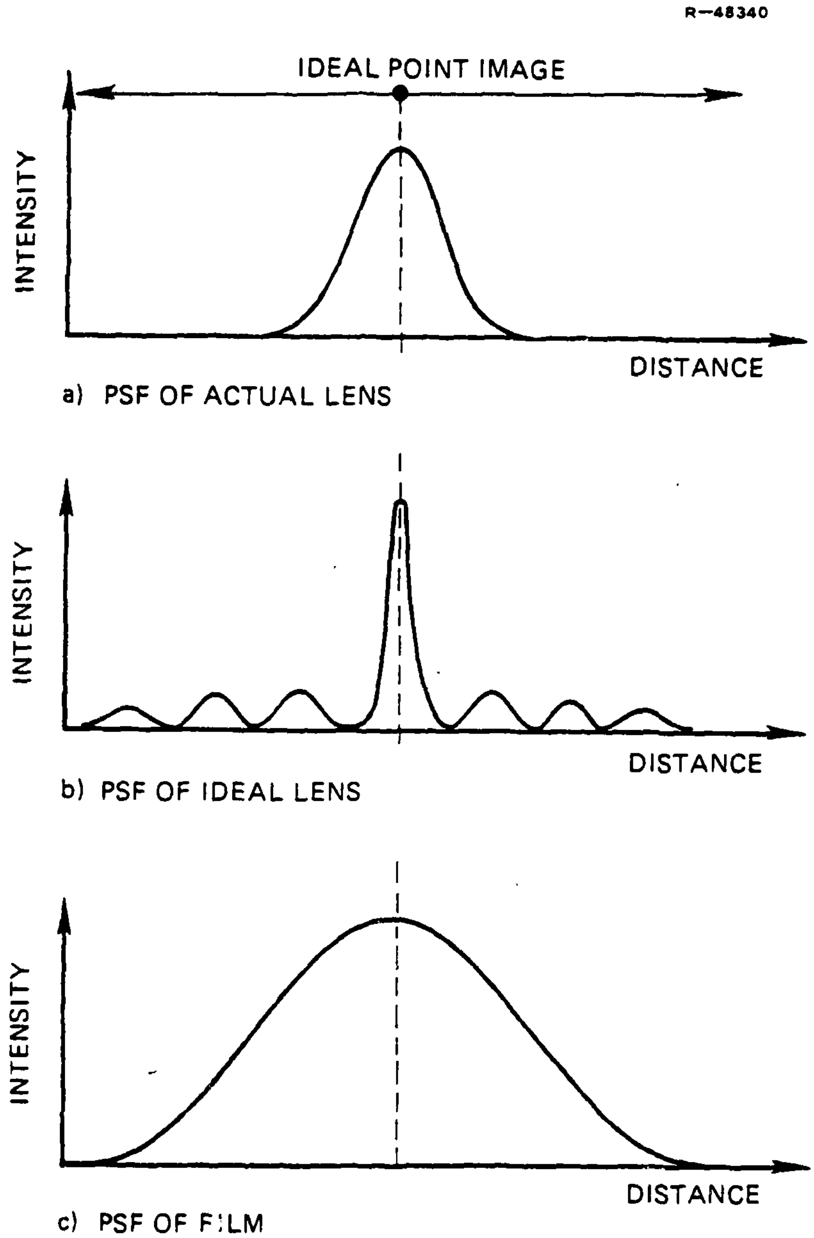


Figure 3.2-23 Examples of Point Spread Functions

- The distribution of light in the object, o(x)
- The PSF of the imaging system, s(x).

by use of the convolution integral

$$i(x) = \int_{-\infty}^{\infty} s(y)\, o(x-y)\, dy \qquad (3.2\text{-}18)$$

where i(x) represents the distribution of light in the image. As shown in Fig. 3.2-24, the convolution integral expresses the fact that the image illumination at any point is a weighted average of the object illuminations in the vicinity of that point, with the weighting given by the PSF. For imaging systems of high quality, the PSF is highly localized; hence the convolution integral (Eq. 3.2-18), although defined theoretically between the limits $-\infty$ and $+\infty$, could for practical purposes be calculated with finite limits -d to +d (Fig. 3.2-24).

At this point comes the key step in modern methods of describing imaging system performance: the application of Fourier analysis methods to the convolution formula (Eq. 3.2-18), in order to characterize system performance in terms of frequency response (frequency domain analysis). The Fourier transforms of the object illumination function, o(x), the PSF, s(x), and the resulting image illumination function i(x), are defined as follows:

$$O(\omega) = \int_{-\infty}^{\infty} o(x) e^{-i\omega x}\, dx \qquad (3.2\text{-}19)$$

$$S(\omega) = \int_{-\infty}^{\infty} s(x) e^{-i\omega x}\, dx \qquad (3.2\text{-}20)$$

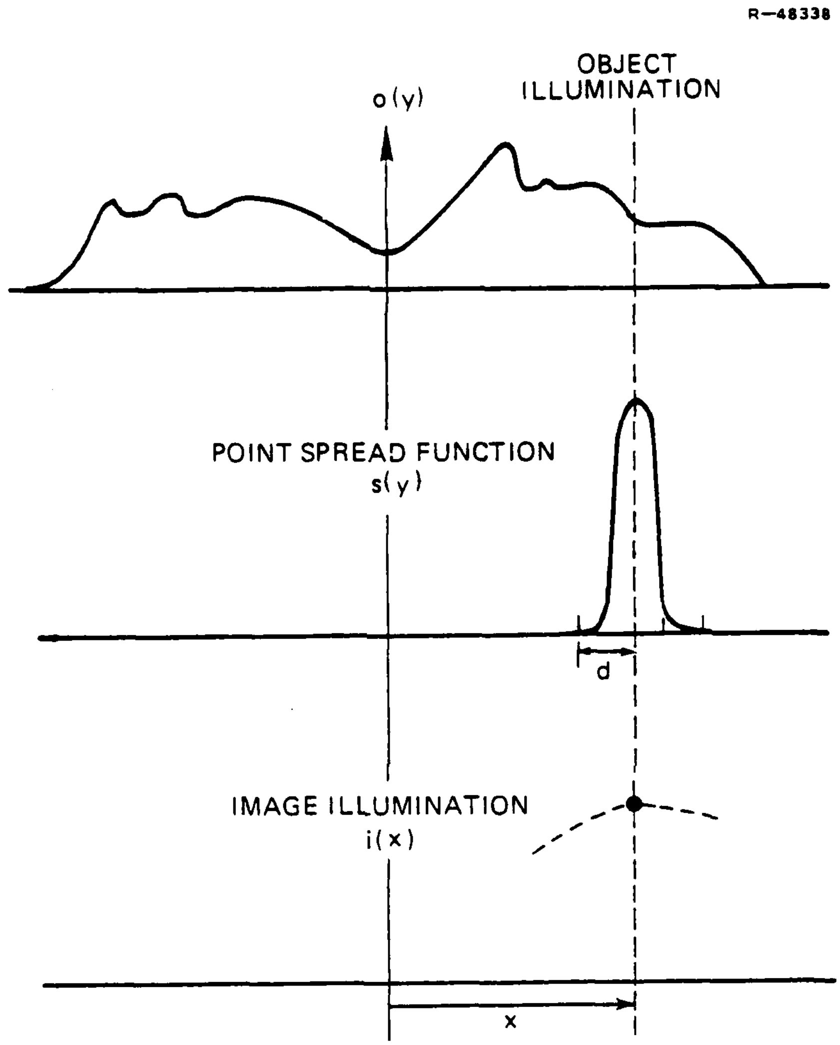


Figure 3.2-24 Concept of the Convolution Integral

$$I(\omega) = \int_{-\infty}^{\infty} i(x)e^{-i\omega x}\,dx \tag{3.2-21}$$

where ω is the frequency variable.

As a consequence of the convolution theorem applied to Eq. (3.2-18), the Fourier transform of the image illumination, $I(\omega)$, can be expressed in the simple form

$$I(\omega) = S(\omega)\ O(\omega) \tag{3.2-22}$$

describing the output (image) in terms of the input (object) and the characteristics of the system. It is customary to normalize the transform of the PSF as follows

$$OTF(\omega) = \frac{S(\omega)}{\int_{-\infty}^{\infty} s(x)\ dx} \tag{3.2-23}$$

to produce the optical transfer function (OTF) -- the normalized Fourier transform of the point spread function.

The OTF as defined by Eq. (3.2-23) is, in general, a complex function. Its real and imaginary parts characterize the amplitude response and the phase shift of the system. Except in the analysis of systems using lasers (or other sources of coherent light), the phase shift information is unimportant, and it suffices to use the absolute value of the OTF to describe the imaging system. The resulting real function is the modulation transfer function (MTF):

$$MTF(\omega) = \left| OTF(\omega) \right| \tag{3.2-24}$$

If the PSF, $s(x)$, is an even function, then its transform, $S(\omega)$, is automatically real and the absolute value step is unnecessary. This is often the case with optical systems of high quality. Also, the spatial frequency, f, is often used in place of the angular frequency, ω:

$$f = \frac{\omega}{2\pi} \tag{3.2-25}$$

Thus the MTF of an imaging system with a symmetric (even) PSF is often defined as

$$\mathrm{MTF}(f) = \frac{\int_{-\infty}^{\infty} s(x) \cos 2\pi f x \, dx}{\int_{-\infty}^{\infty} s(x) \, dx} \tag{3.2-26}$$

Equation (3.2-26) is not only a theoretical definition, but shows how MTFs are determined in practice:

- Illuminate the system with a point source
- Measure the image illumination (PSF)
- Evaluate the integrals of Eq. (3.2-26).

It is useful to contrast this practical procedure with the basic concepts expressed in Eqs. (3.2-13) through (3.2-17), the implementation of which would require:

- The construction of a bar pattern test object for every frequency f
- The measurement of the image modulation to get the MTF at that frequency.

The numerical integration operation embodied in the practical evaluation of Eq. (3.2-26) is usually a simpler process.

CHAPTER THREE
IMAGING OPERATIONS

Once an image has been formed, by the use of one or another of the imaging systems discussed in Chapter Two, it may be necessary to operate on the image (i.e., to modify or transform it) in order to extract as much information as possible or to make it more suitable for viewing, analysis, and interpretation. Three major aspects of imaging operations are treated in this chapter:

- Altering the geometry of the image (to correct distortions, for example) -- Section 3.3.1
- Scanning, sampling, and digitizing the image -- Sections 3.3.2 and 3.3.3
- Enhancing the image -- Section 3.3.4.

Geometric alterations to an image may be desired for a variety of reasons. In the case of an aerial or satellite photograph, for example, corrections are necessary to compensate for the effects of lens distortion, height displacement distortion, and possible deviation of the line of sight from the true vertical (review Section 1.2.3 of Unit One). Even with such corrections, the photograph cannot be superimposed on a standard map unless a transformation is applied to relate the photographic image to the particular map projection (Section 1.2.6 of Unit One) being used.

Much modern image processing is done by computer programs operating on a discrete representation of the image. In

terms of an original photographic image, the creation of a discrete representation involves two processes:

- Scanning and sampling -- in which the original continuous image is segmented into a number of discrete picture elements (pixels), and with a gray-scale value representing an average over the area of the original image represented by the pixel

- Digitization -- in which the continuous gray-scale range is represented by a relatively small number (usually a power of two) of discrete steps.

For some imaging systems, one or both of these stages may already have been carried out while the image was being formed. Otherwise, the required image dissection and analog-to-digital conversion must be accomplished by suitable interface equipment prior to computer processing of the image.

Image enhancement includes a wide variety of techniques for improving the visual information content of a picture and for detecting various features of interest that may be difficult to discern in the original images. Some of these techniques include

- Deblurring (sharpening a picture that is blurred because of defocusing or relative object-imager motion)
- Enhancement of contrast
- Improvement of signal to noise ratio
- Enhancement of edges and other features.

Image enhancement is the subject of Section 3.3.4.

3.3.1 Rectification and Orthoimagery

The first step in image processing is rectification. This consists of eliminating the variations in scale* within the image caused by tilt of the camera during the exposure process.

Figure 3.3-1 represents the camera-terrain geometry at the time of exposure. The point O is the perspective center, or point of perspective, of the photograph (i.e., the camera location). The optical axis, OP, is perpendicular to the photographic plane. A vertical line intersecting the photographic plane and passing through the perspective center marks the nadir point, N. The tilt angle is denoted α. Rectification involves elimination of the scale variation due to α. In the special case of perfectly flat terrain, rectification is equivalent to moving the perspective center to make O, P, and N collinear; in this case, the rectified image is the image that would have been obtained if the camera had been directly overhead and pointing straight down at the time of exposure.

The process is illustrated in Figure 3.3-2. The original photograph (left side) shows severe scale distortion, principally along the horizontal axis, because of the camera tilt angle. The rectified image, obtained by a geometric transformation of the original image, is shown at the right. The uniformity of scale makes the image at the right much more convenient to work with in cartographic data extraction operations than the original image.

---

*Ratio of distance in the photograph to distance on the ground. Ideally, the scale is constant all over the photograph, but camera tilt and certain other effects can introduce undesirable distortions.

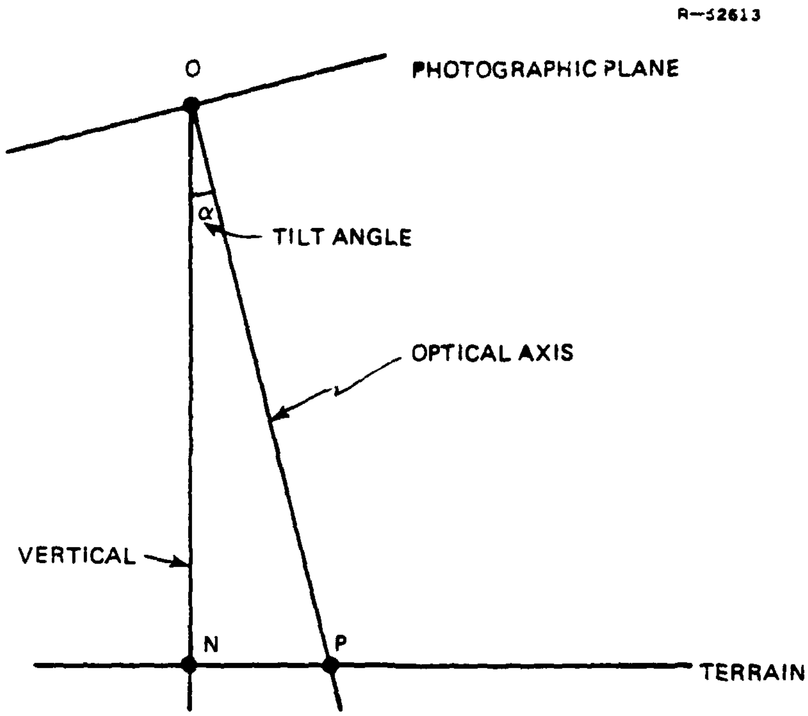


Figure 3.3-1 General Representation of Camera-Terrain Geometry

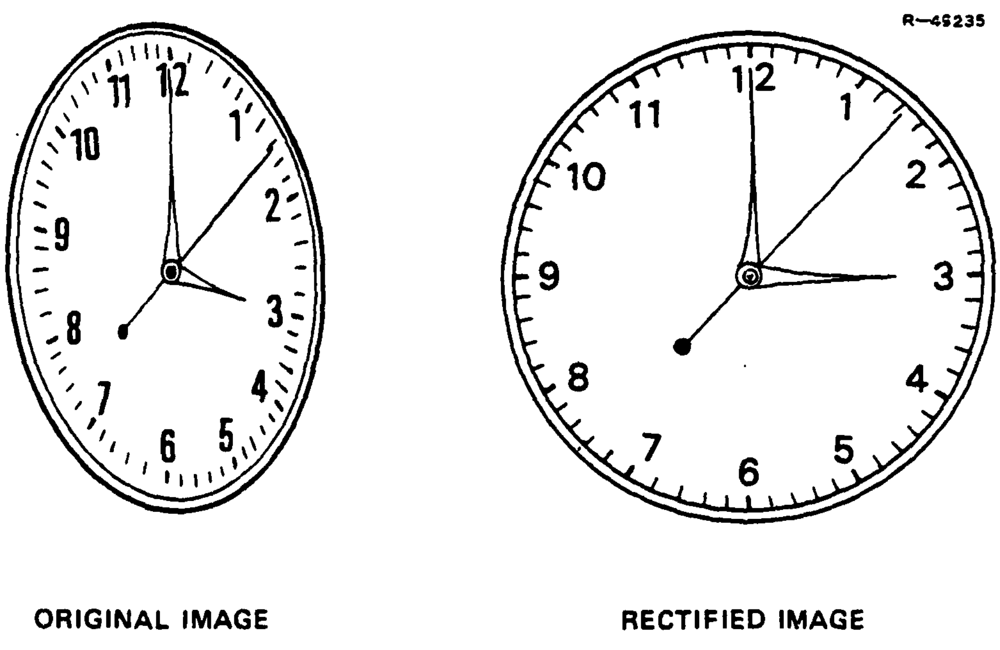


Figure 3.3-2 Rectification Example

The foregoing discussion assumes that the photograph is made instantaneously. In this case, rectification is a straightforward geometric transformation, which can be implemented by optical reprojection techniques. More complicated cameras exist, however, in which the exposure takes place over an extended period of time.* Significant relative motion between the camera and the imaged terrain can then take place during the exposure process, leading to complex distortion patterns. Finally, the relative geometry between the camera and imaged terrain may not be known accurately. All these factors can lead to additional levels of complication in rectification operations.

Another image processing operation may be required before data extraction takes place. This is orthographic projection, and involves removing residual distortions in the imaging due to variations in the terrain height (terrain relief) within the imaged scene. Terrain height variations cause scale changes because of the varying distance between the camera and points in the image. (This is the effect that causes the top of a building to appear larger than its base in a photograph taken from close overhead.) Orthographic projection requires additional processing, using previously compiled elevation data, to remove this distortion. The result is known as an orthophoto.

3.3.2 Sampling and Digitizers

Although actual imagery is essentially continuous in nature (i.e., a continuous spatial energy distribution defined on continuous variables (x,y) at time t, and wavelength λ), it is generally necessary to represent an image as a finite number

---

*For example, in panoramic cameras with a moving-slit shutter.

of discrete picture elements (pixels). This is done by <u>sampling</u> the image. It is also necessary to express the energy level at any point of the image (or for any pixel) as one of a finite number of discrete energy levels (<u>quantization</u>). Sampling and quantization are required in order to represent the image in digital form for storage and processing by computer. In this section, the sampling process and practical hardware to implement the sampling operation are considered.

<u>Sampling Process</u>(†). The process of sampling and quantization may be thought of as a method for transforming a continuous image field $f(x,y)$ to a set of discrete values, $f_k(i\Delta x, j\Delta y)$, given by the integers i,j,k. This may be accomplished as a two-step process in which the continuous field is sampled at the points $x = i\Delta x$, $y = j\Delta y$ and then the result, $f(i\Delta x, j\Delta y)$, is quantized to one of k discrete levels.

In a perfect image sampling system, spatial samples of the ideal image would, in effect, be obtained by multiplying the ideal image by a spatial sampling function

$$s(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \delta(x-i\Delta x,\ y-j\Delta y) \tag{3.3-1}$$

composed of an infinite array of Dirac delta functions arranged in a grid with spacing $(\Delta x, \Delta y)$ as shown in Fig. 3.3-3. The sampled image is then represented as

$$f(i\Delta x, j\Delta y) = f(x,y)s(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i\Delta x, j\Delta y)\ \delta(x-i\Delta x, y-j\Delta y) \tag{3.3-2}$$

(†)This section contains material at a more advanced level than the rest of the text.

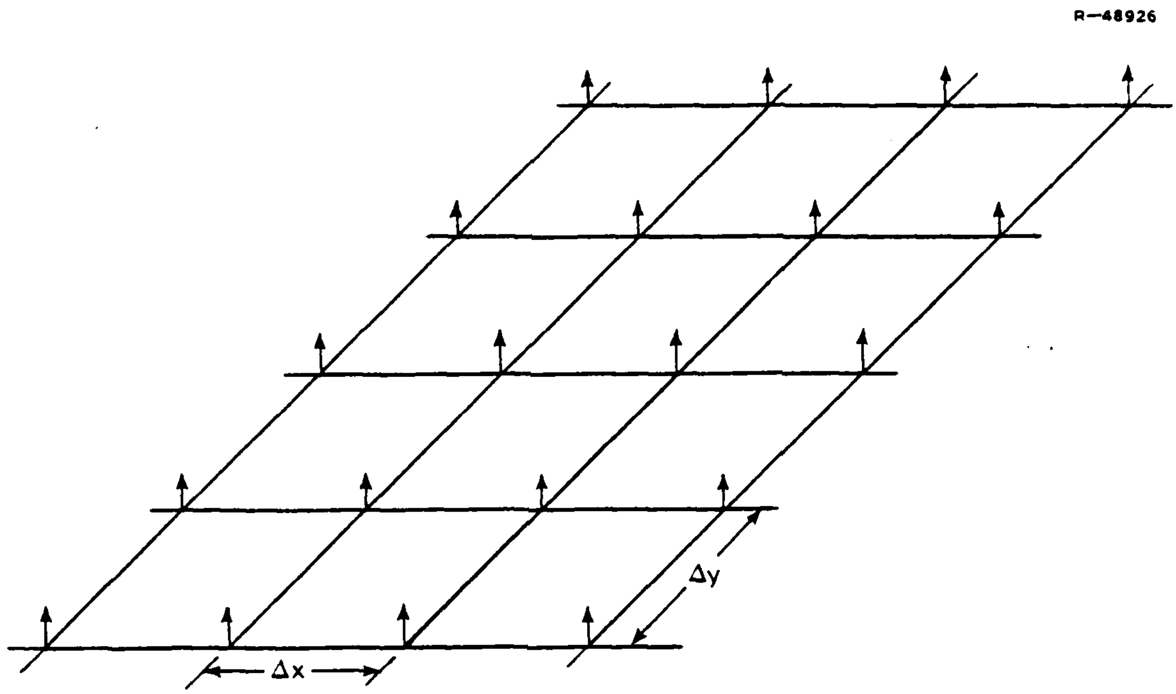


Figure 3.3-3 Dirac Delta Function Sampling Array

The function f(x,y) may be brought inside the summation and evaluated only at the sample points $(i\Delta x, j\Delta y)$. It is convenient, for purposes of analysis, to consider the spatial frequency domain representation (ω_x, ω_y) of the sampled image, obtained by taking the continuous two-dimensional Fourier transform of the sampled image. Thus

$$F_s(\omega_x,\omega_y) = \int_{-\infty}^{\infty} f(i\Delta x, j\Delta y))\exp\{\sqrt{-1}(\omega_x x + \omega_y y)\}\ dx\ dy \tag{3.3-3}$$

The Fourier transform of the sampled image can be expressed as the convolution of the Fourier transforms of the ideal image $F(\omega_x,\omega_y)$ and the sampling function $S(\omega_x,\omega_y)$, as expressed by

$$F_s(\omega_x,\omega_y) = F(\omega_x,\omega_y) * S(\omega_x,\omega_y) \qquad (3.3\text{-}4)$$

where * represents the two-dimensional convolution operation.

Convolution in one dimension has been introduced in Section 3.2.6. In two dimensions it is defined as

$$F_s(\omega_x,\omega_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\omega_x-\sigma_1,\omega_y-\sigma_2)\ S(\sigma_1,\sigma_2)\ d\sigma_1\ d\sigma_2 \qquad (3.3\text{-}5)$$

and abbreviated as in Eq. (3.3-4). The two-dimensional Fourier transform of the spatial sampling function is an infinite array of Dirac delta functions in the spatial frequency domain

$$S(\omega_x,\omega_y) = \frac{1}{\Delta x\ \Delta y}\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} \delta(\omega_x-i\omega_{xs},\omega_y-j\omega_{ys}) \qquad (3.3\text{-}6)$$

where $\omega_{xs} = 2\pi/\Delta x$ and $\omega_{ys} = 2\pi/\Delta y$ are the sampling frequencies. It is assumed that the spectrum of the ideal image is bandlimited to some bounds such that $F(\omega_x,\omega_y) = 0$ for $\omega_x > \omega_{xc}$ and $\omega_y > \omega_{yc}$. That is, there is no signal content in the image above the cutoff frequencies ω_{xc} and ω_{yc}.

Performing the convolution of Eq. (3.3-4) yields

$$F_s(\omega_x,\omega_y) = \frac{1}{\Delta x\ \Delta y}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\omega_x-\alpha,\omega_y-\beta)$$

$$\cdot \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} \delta(\alpha-i\omega_{xs},\beta-j\omega_{ys})d\alpha\ d\beta \qquad (3.3\text{-}7)$$

where α and β are new variables of integration. Upon changing the order of summation and integration and invoking the <u>sifting</u>

property* of the delta function, the sampled image spectrum becomes

$$F_s(\omega_x,\omega_y) = \frac{1}{\Delta x\ \Delta y} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} F(\omega_x - i\omega_{xs}, \omega_y - j\omega_{ys}) \qquad (3.3\text{-}8)$$

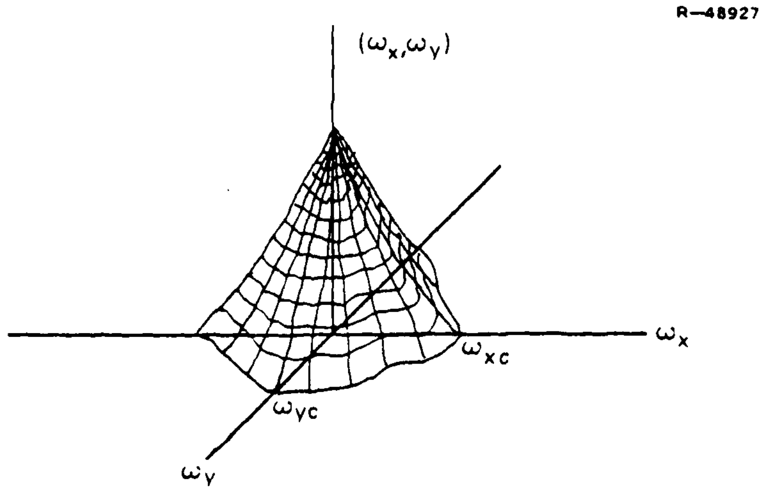
As can be seen from Fig. 3.3-4, the spectrum of the sampled image consists of the spectrum of the ideal image infinitely repeated over the frequency plane in a grid with a frequency spacing of $(2\pi/\Delta x, 2\pi/\Delta y)$. It should be noted that if Δx and Δy are not sufficiently small as compared with the spatial frequency limits of $F(\omega_x,\omega_y)$ -- i.e., unless

$$\frac{1}{\Delta x} \leq 2\omega_{xc} \quad \text{and} \quad \frac{1}{\Delta y} \leq 2\omega_{yc} \qquad (3.3\text{-}9)$$

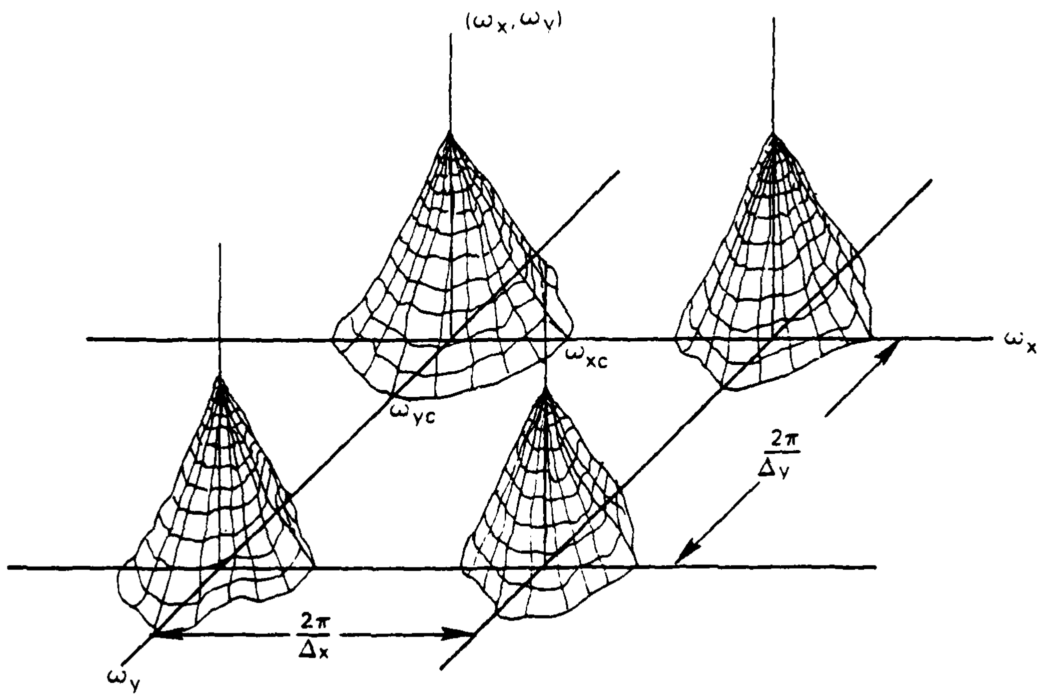
then the individual spectra will overlap, and it will not be possible to reconstruct the original image from the samples. This Fourier frequency overlap is known as aliasing. To avoid its effects, the conditions in Eq. 3.3-9 must be satisfied. This can be paraphrased as requiring that the sampling rate (i.e., the inverse of Δx and Δy) must be at least twice the highest frequency in the signal to be sampled (i.e., $2\omega_{xc}$ and $2\omega_{yc}$), which is sometimes known as the Nyquist frequency.

In order to reconstruct the original image $f(x,y)$ from its sampled version, $f(i\Delta x, j\Delta y)$, a low-pass filter is used. The low-pass filter eliminates all of the spectral peaks except the one at the origin -- i.e., the $i=j=0$ term in Eq. 3.3-8 is retained. If a rectangular low-pass filter, $h(x,y)$, is used, whose Fourier transform value is one in the frequency

*That is, $\int_{-\infty}^{\infty} f(x)\ \delta(x-y)dy = f(y)$.



a) ORIGINAL IMAGE FOURIER SPECTRA



b) SAMPLED IMAGE FOURIER SPECTRA

Figure 3.3-4 Spectra of a Typical Sampled Image

range ($\pm\, \omega_{xc}$, $\pm\, \omega_{yc}$) and <u>zero</u> elsewhere, f(x,y) is given by the two-dimensional function

$$f(x,y) = h(x,y) * f_s(i\Delta x, j\Delta y)$$

$$= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f_s(i\Delta x, j\Delta y) \operatorname{sinc}[2\omega_{xc}(x-i\Delta x)] \operatorname{sinc}[2\omega_{yc}(y-j\Delta y)] \quad (3.3\text{-}10)$$

where the sinc function is defined by

$$\operatorname{sinc} x = \frac{\sin \pi x}{\pi x} \quad (3.3\text{-}11)$$

Filters with other than the rectangular low-pass characteristic can be used. The use of different filters results in different versions of the reconstructed function.

The sampling process described above is clearly idealized; in actual applications there are many additional constraints. The effects of practical limits are

- The sampling system is never a Dirac delta function sampler; consequently, the spectrum of the sampled image is degraded by the transfer function resulting from finite width of the sample spot.
- The reconstruction and display system cannot be a true sinc function response since this implies negative light and a display spot of infinite extent (because of the negative portions and infinite extent of the sinc function). Realistic display systems use positive light and finite spots; consequently, there is never perfect separation of the zero-order spectral peak from higher-ordered, repeat spectra of the data.
- The effects of spectrum degradation resulting from finite sampling and display spots can be computed and corrected by an image restoration process. If the original data are reasonable, the result

of these operations is a displayed digital image that has better resolution and fine structure detail.

- The lack of perfect separation of the zero-order spectral peak from higher repeat spectra is seldom a problem in displays; it is usually visible only in special cases. An example would be a periodic structure with frequencies near the Nyquist sampling limit.

The reader interested in practical applications is encouraged to consult references given in the reading list at the end of Unit Three.

Once the image $f(x,y)$ has been sampled, the sampled function $f(i\Delta x, j\Delta y)$ must be quantized for storage in digital form. This means that the continuous range of values possible for $f(i\Delta x, j\Delta y)$ must be divided into k discrete bins or regions that are represented by numbers in the computer. The placement of the decision regions on the f axis is usually dictated by a desire to minimize quantization noise based on various fidelity measures. The interested reader is referred to the reading list for more detailed discussion. Many ingenious quantizing schemes have been developed. For example, a nonuniform quantizer using logarithmic compression provides excellent results in minimizing mean-squared quantization error and distortion. The scheme works well for signals with a wide variety of statistical properties.

The quantized sample of $f(i\Delta x, j\Delta y)$ is denoted as $f_k(i\Delta x, j\Delta y)$ and is usually referred to as a pixel (picture element). A typical image has values for i and j such that $0 < i,j < 512$. As a result, such an image has over a quarter million pixels. Image sizes of hundreds of millions of pixels are possible from some of the high-resolution sensors on weather

and special-purpose satellites. Values for k, the number of quantization levels, have been determined largely by psychovisual studies which indicate that the maximum number of levels discernable by the human visual system is about 256. Thus 256 levels (shades of gray), requiring 8 bits of information to represent ($2^8 = 256$), have become the norm in most image processing applications. However, in some high-resolution applications, quantization levels ranging from one bit to 16 bits are used.

3.3.3 Scanning Operations and Equipment

Types of Scanners - Practical methods for digitizing continuous imagery from sources such as photography fall into four categories depending on the sensor selected:

- Television camera
- Scanning microdensitometer
- Photodiode array
- Laser.

Each system provides a different set of tradeoffs in terms of accuracy, speed, resolution, and ease of use.

Television camera digitizers are available in two basic formats, raster scan and random scan. The raster scan format (described in Section 3.2.2) consists of rapidly scanning the electron beam across the face of the tube from side to side, with each successive line scan occurring below its predecessor. The most common format in the United States uses 525 lines, but other formats vary from several hundred to several thousand lines for special CCTV (Closed Circuit Television) applications. The raster scan generates data at a very high

speed. For example, even the least expensive units generate data at a rate of 10^5 points per second. The information is digitized to 64 levels of gray and results in data rates of 6×10^6 bits/second. Until recently, this rate severely taxed the capability of analog-to-digital converters and computers used to process the image data. However, some of the modern image display systems are capable of digitizing and storing 80×10^6 bit/second television signals for several frames of data.* A number of techniques have been developed, on the other hand, to permit the transmission of scanned imagery at lower data rates. These include slow scan (e.g., one frame/ sec instead of 60) and subsampling (e.g., taking one sample per line and changing the sample point each frame). These techniques essentially increase the time to digitize a 525 line frame from about a thirtieth of a second to as much as 30 sec.

Random scan techniques move the beam in any desired pattern over the face of the tube under X-Y control from the computer. Although the data rate is considerably slower than in the raster scan technique, this limitation can often be overcome by selectively moving the beam to digitize only the regions of interest. Figure 3.3-5 shows a typical television camera intended for use as a sensor for computer digitizing.

In summary, television camera digitizers provide high speed but relatively low resolution and accuracy. They are physically compact with considerable optical and mechanical flexibility to handle a wide variety of image types and formats.

The scanning microdensitometer consists of a light source, a photodetector, and a precise means for accurately

*One frame is a full picture.

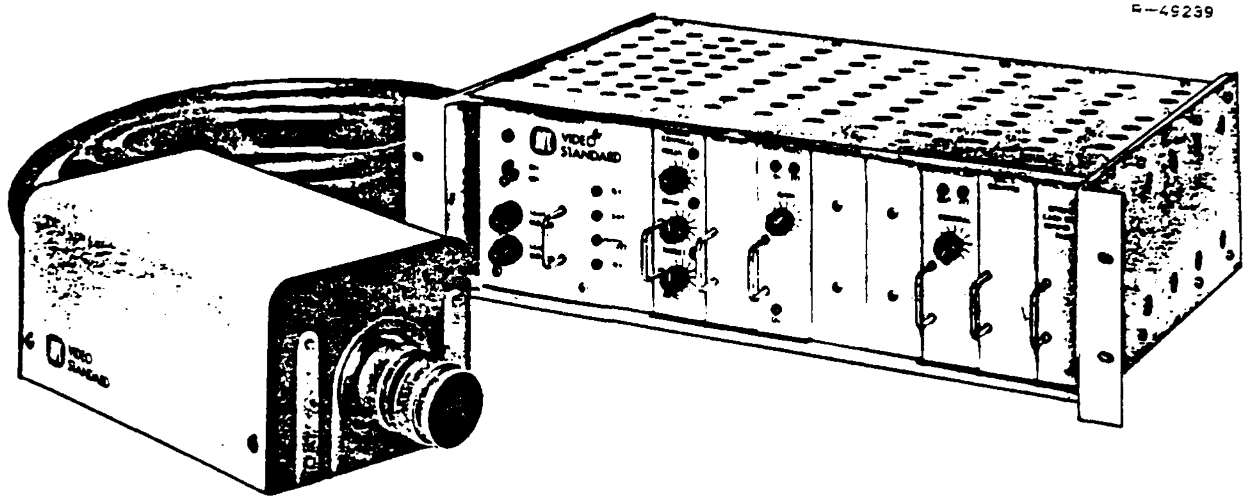
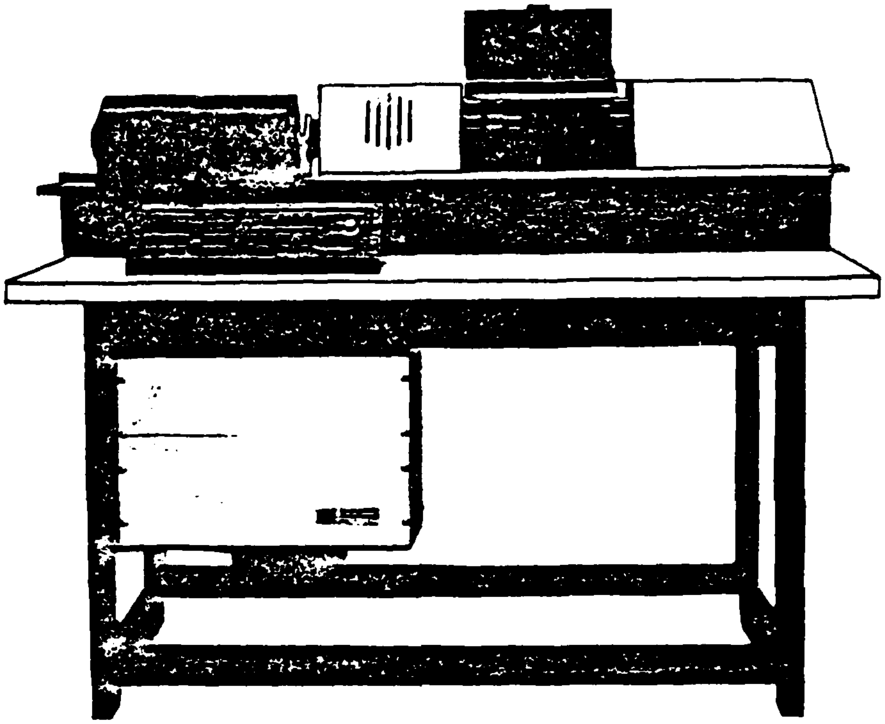


Figure 3.3-5 Typical Television Camera System Intended for Computer Digitizing of Image Data

moving the film between the source and detector. The light focuses onto a small, usually square, area of the picture. The amount of light transmitted through this area of film is then compared to light transmitted with no image present.

Two basic designs, rotating drum and flatbed, find common use as high precision digitizing and recording instruments. Rotating drum systems combine moderate scan and plot speeds with excellent characteristics in image resolution, density discrimination, pixel accuracy, chromatic fidelity, and the ability to digitize and plot very large image formats. Flatbed systems, while not as fast as the rotating drum, offer greater accuracy, precision, and resolution. In the flatbed system, a precision X-Y table transports the film. Larger flatbed systems are mounted on granite blocks to provide thermal isolation of the optics and vibration-free operation. Figure 3.3-6 shows typical examples of both types of microdensitometers

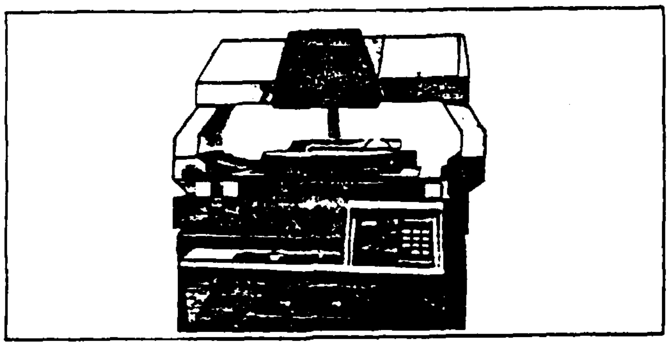


(a) Typical Rotating Drum Scanning Microdensitometer

Figure 3.3-6 Scanning Microdensitometers for Digitizing Data for Input to Computer Systems

Scanning microdensitometers provide a high degree of accuracy and resolution, but have quite slow scanning speeds. They tend to be fairly bulky and costly, and require considerable maintenance and adjustment to keep the high-accuracy mechanical positioning unit in adjustment. They are generally the first choice when accuracy is the paramount performance parameter.

The <u>photodiode array</u> or <u>array scanner</u> combines some of the features of the television camera and the microdensitometer. Essentially, the array is a sequence of photosen-



(b) Typical Flatbed Scanning Microdensitometer

Figure 3.3-6 Scanning Microdensitometers for Digitizing Data for Input to Computer Systems

sitive devices as shown in Fig. 3.3-7, coupled to an electronic shift register to read out the data from the photodiodes. A typical sensor has 1024 elements of size 25 μm by 2.5 mm with 25 μm spacing between elements (i.e., in Fig. 3.3-7, N=1024, a=25 μm, b=50 μm, c=2.5 mm). The arrays are available in a wide variety of sizes and shapes in both one and two dimensional arrays, as shown in Fig. 3.3-8.

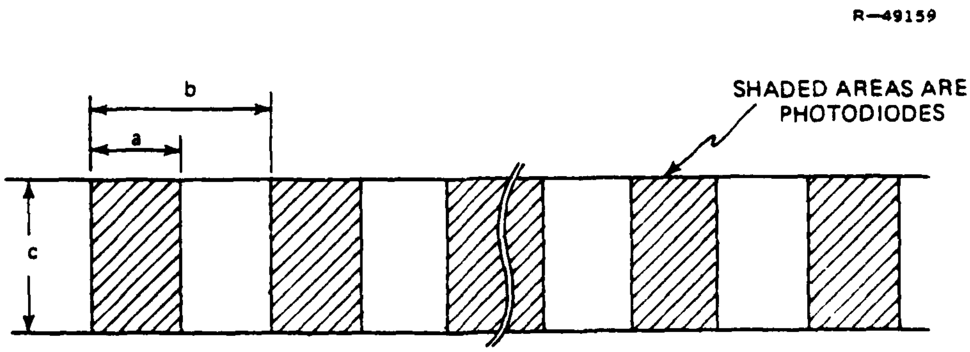
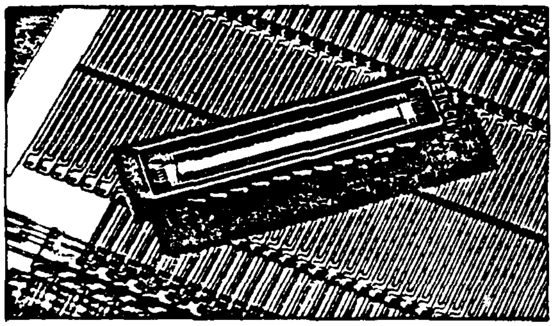
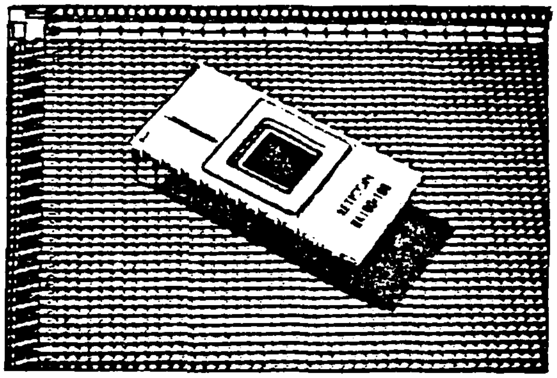


Figure 3.3-7 Typical Sensor Geometry of N-Element Linear Array

R-49241



a) Photograph of 1024 element linear array with enlarged view of silicon chip in background. One mil wide sensing aperture runs down center of chip.



b) Photograph of a 100x100 matrix array with enlarged view of silicon chip in background.

Figure 3.3-8 Typical Examples of a Linear and a Two-Dimensional Photodiode Array

The arrays may be used in television applications by replacing the vidicon tube with the photodiode array or by moving the film and/or array to do the scanning. Figure 3.3-9

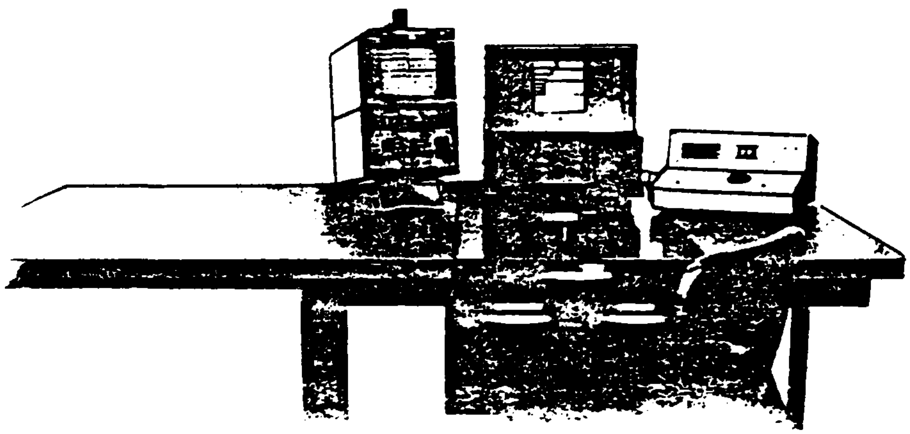


Figure 3.3-9 Linear Array Scanner for Digitizing Film Transparencies

shows an example of a moving array scanner. It scans images containing 1024 by 1024 output pixels (of 8 bits each) from film transparencies in about 17 sec.

Photodiode arrays represent a compromise between the television camera and microdensitometer, providing nearly the same accuracy and resolution as the microdensitometer, but with several orders of magnitude increase in speed. Their cost tends to be moderate, and they can be used in compact designs like the television camera digitizers or in more elaborate scanning systems based on microdensitometer moving platforms.

Laser scanners use a laser light source which is rapidly scanned across the film surface by a combination of rotating mirrors and electro-acoustical deflectors. The monochromatic, coherent light provided by the laser allows a very small

spot size and high resolution. Because the beam may be deflected rapidly, very high digitization rates are also possible. The major drawbacks are the elaborate optical system, the difficult alignment problems, and the relatively high cost of this type of scanner.

Laser scanners are used in situations where maximum scanning speeds must be coupled with moderate to high accuracy and resolution. They are very bulky and require precise alignment and calibration, but offer scanning speeds which are usually limited only by the ability of the computer registers receiving their output to keep up.

The scanner types described in this section are very different. Because their technologies have evolved in dissimilar contexts, there are no commonly accepted criteria for resolution, dynamic range, or signal-to-noise ratio among the various manufacturers. However, Table 3.3-1 provides a summary of some of the more important scanner characteristics for typical hardware, adjusted to provide common definitions where possible.

3.3.4 Image Enhancement

Applications - The image enhancement techniques described in this section are preprocessing operations which make it easier to extract information from the imagery. These techniques range from improving the quality of images degraded by some unknown process, to methods which deliberately distort the imagery in a manner which complements the capabilities of the human visual system. Also included are techniques which serve as effective prefilters to support highly automated pattern recognition operations.

TABLE 3.3-1

COMPARISON OF THE PERFORMANCE OF SEVERAL SCANNER TYPES

T-3978

SCANNER TYPE	SPEED IN PIXELS/SEC	RESOLUTION	ACCURACY	IMAGE FORMAT SIZE	OUTPUT LEVELS	NO. OF DIMENSIONS CHARACTERIZED
Raster Scan T.V. Camera	10 million	Depends on lens magnification 512 to 512 pixels	±2 percent of full scan	Depends on lens magnification	6 to 8* bits	2 or 3
Random Scan T.V. Camera	10 to 100 thousand	Depends on lens magnification, 1024 by 1024 points	±1 to 2 percent of full scan	Depends on lens magnification	4 to 8* bits	2 or 3
Rotating Drum Microdensitometer	30 to 100 thousand	10 to 50 μm spot sizes	2 μm/cm	Up to about 20 by 20 in	8 to 10 bits	3
Flatbed Microdensitometer	1 to 10 thousand	1 to 200 μm spot size	1 μm/cm	Up to about 10 by 10 in	8 to 10 bits	4
Photodiode Array	0.1 to 1 million	Depends on lens magnification 1024 by 1024 pixels 1 to 16 μm spot size	1 to 10 um/cm	Up to about 10 by 10 in	8 bits	3
Laser	0.2 to 10 million	1 to 50 μm spot size	2 to 5 μm/cm	Up to about 20 by 20 in	8 to 10 bits	3

*With averaging or integration.

Practical applications of image enhancement have been extremely widespread, covering many diverse areas. Among the most successful have been:

- Deblurring of imagery from optical devices to remove the effects of atmospheric distortion, vehicle motion, or film grain
- Increasing the contrast and/or intensity of special imagery such as x-rays and sonagraphs
- Reduction of noise and restoration of missing data values in transmitted images
- Mapping black and white images into color (pseudocoloring) for enhancing the detectability of small gray scale differences
- Sharpening boundaries for subsequent processing and data extraction (e.g., fingerprints, roads, area calculations).

Because there is no universally accepted agreement concerning the definition of image quality, evaluation of the merits of most image enhancement techniques is something of an ad hoc exercise. However, on the basis of experiments with both human viewers and automatic algorithms, a number of general enhancement techniques have proved to be useful in a wide variety of applications. These techniques can be divided into three major categories:

- Noise reduction
- Gray scale manipulation
- Edge enhancement and sharpening.

In the next sections, representative techniques in each of these categories are described in more detail.

Noise Suppression -- The effects of noise on the imagery diminish the ability of human operators to perceive fine detail, and can cause difficulty for automatic algorithms that are based on differentiation, correlation, or contrast observation. Suppression or removal of the noise is of interest in situations such as the following:

- Correction of noisy and/or dropped pixels or lines in the image record caused by transmission noise or recording media errors
- Minimization of the effects of compression-induced noise in subsequent processing operations
- Reduction of the effects of errors introduced by resamp ing and other photogrammetric processing operations
- Suppression of undesirable effects brought about by other enhancement techniques such as contrast stretching.

Table 3.3-2 describes a number of potential techniques for reducing noise effects in the input imagery. The actual approach to implementing these techniques involves estimating the characteristics of the noise likely to be encountered, and applying various noise suppression algorithms. An appropriate algorithm is selected by either observing the results or measuring the impact of the suppression operation on the efficiency of subsequent image processing activities.

Gray Scale Manipulation -- Imagery that is of practical interest to users often has distorted gray scale distributions caused by conditions such as lens flare, haze, and film or sensor nonlinearities. Even imagery with undistorted gray scales can be far from optimal for photo-interpretation purposes. For example, low-contrast regions of a scene may actually contain

TABLE 3.3-2

TYPICAL NOISE REDUCTION TECHNIQUES

APPROACH	DESCRIPTION
Outlier Removal (e.g., median filtering)	Identification and removal of irregular pixels
Smoothing	Averaging out random noise without substantially blurring the imagery
Clipping	Utilization of a priori knowledge of bounds on the true illumination levels to eliminate noise spikes
Interpolation	Fitting smooth functions to noisy data
Logarithmic Filtering	Special preprocessing to reduce multiplicative noise effects
Spatial Filtering (e.g., Wiener filtering)	Filtering using a priori statistical knowledge of the noise field

only a few gray levels, while a typical electronic display is capable of portraying up to 256 gray shades at once. Stretching of the gray levels in the scene to match the capabilities of the display (i.e., mapping the darkest gray level into black and the lightest gray level into white, with a commensurate scale adjustment in between) greatly enhances the apparent contrast in the scene, thereby improving resolution. There are many other applications of gray scale manipulation as an image enhancement tool.

In more general terms, gray scale manipulation usually involves simple, but powerful, nonlinear point processes which map the radiometric intensity of each pixel into a new

value independent of its neighbors. These techniques are generally implemented in the computer by table look-ups, in which the true radiometric intensity is used as a pointer to determine the adjusted value. The tables may be changed dynamically, or from region to region within an image. Pseudocolor mapping is a further generalization of this concept. Table 3.3-3 contains descriptions of some typical gray scale manipulation techniques.

Edge Enhancement and Sharpening - Imagery blurring and edge smearing can occur for a number of reasons, including uncompensated platform motion, imperfectly focused optics, atmospheric turbulence, and other unintentional low-pass filtering processes. In addition, low-contrast scenes and the basic physics of the optics (resolution limits) tend to produce blurred or fuzzy edges in the image. The techniques described in Table 3.3-4 may be used to deblur the imagery, or alternatively to enhance the edges (a closely related process). Naturally, the exact technique to use in any specific application depends on the quality of the original imagery and the amount of sharpening desired.

Implementation Considerations. In addition to the appropriateness of the enhancement techniques just described, two other considerations must usually be taken into account -- computation speed and generation of artifacts. The computation speeds for these algorithms on a general-purpose computer range from a few seconds to tens of minutes. It has been found from experimental studies that techniques taking more than 10 to 20 sec to execute are simply not used by photo-interpreters in real-time environments. Even in non-real-time applications, the longer algorithms tend to use inordinate amounts of relatively expensive computer time, particularly for large image sizes.

TABLE 3.3-3

TYPICAL GRAY SCALE MODIFICATION TECHNIQUES

APPROACH	DESCRIPTION
Histogram Manipulation	Rearrangement of an image's gray scales to provide an alternate (e.g., flat) histogram. This provides more detail by redefining the gray levels in proportion to their number of occurrences.
Retinal Model Contrast Control	Local area brightness and contrast control
Contrast Stretching	Emphasis of particular gray scale values at the expense of others
Bit Slicing	Use of only certain bits of the binary representation of the gray scale value (useful in contour determination)
Gamma Correction	Inversion of film and display nonlinearities
Homomorphic Filtering	Extraction of detail from contrasting parts of the imagery without destroying the overall balance
Gray Level Slicing	Simple technique for providing excellent contouring and edge information
Gray Scale Reversal	Conversion from negative to positive images and vice versa
Pseudocolor Mapping	Mapping of black and white images into variations of color to accentuate small gray scale differences

TABLE 3.3-4

TYPICAL EDGE ENHANCEMENT AND EDGE SHARPENING TECHNIQUES

APPROACH	DESCRIPTION
Spatial Differentiation	Calculation of Laplacian or spatial gradient (usually sensitive to noise)
High-Pass Filtering	Compensation for blurring, which usually weakens high frequencies more than low frequencies
Blind Deconvolution	Use of restoration techniques, but guessing at the form of the degrading transfer function (e.g., Gaussian)
Homomorphic Filtering	Logarithmic transformation, followed by high-pass filtering (sharpens an image and increases detail in high-contrast areas)
Unsharp Masking	Subtraction of the low-frequency portion of an image (also known as DC suppression)
Alpha Processing	Nonlinear transformation which emphasizes phase and edge information
Convolving Mask	Image motion suppression
Edge Operators (e.g., Sobel, Roberts)	Spatial filters designed to detect edges with less noise than ordinary spatial differentiation

In an effort to reduce the computational burden, a whole new generation of special purpose processors has been developed. These include array, parallel, and video-rate display processors. These hardware devices allow many of the more basic enhancement routines, like spatial filtering and

edge detection, to be done in a fraction of a second instead of several minutes. Development of multiple processor architectures and single-chip computers is likely to lead to even more cost-effective implementations of current algorithms and provide increased capabilities for future, more sophisticated techniques.

Note that a number of the enhancement techniques can lead to artifacts in the processed image. These effects range from halos around edges to blurring of certain image areas during noise reduction. If the imagery is to be used for the detection of particular characteristics (e.g., objects in reconnaissance photographs), then it is mandatory that the user be aware of the types of distortions or artifacts that a given enhancement process can give rise to. However, familiarity with these artifacts usually comes with experience. Seasoned observers are normally quite successful in avoiding most artifact-induced problems.

CHAPTER FOUR

ATMOSPHERIC EFFECTS ON IMAGING SENSORS

In reconnaissance and remote sensing of the earth's surface from aircraft or spacecraft, the earth's atmosphere degrades the observations. Some parts of the electromagnetic spectrum are less attenuated than others in passing through the atmosphere. These regions of the spectrum, known as windows, are not only those traditionally used by ground-based astronomy, but also are generally the most useful for long-range sensing of the earth's surface. As different ground-sensing missions have different operational requirements, it is important to consider atmospheric effects on sensor selection.

Figure 3.4-1 shows the atmospheric attenuation in dB/km as a function of wavelength for sea level pressure, clear sky conditions, and moderate humidity (7.5 g/m^3 of water content). These curves represent approximate values of attenuation, and are intended to show overall trends instead of exact numerical values. The commonly used names of the various spectral regions are also shown. Frequencies below 1000 GHz are also illustrated.

In general, clear-air atmospheric attenuation of electromagnetic radiation occurs through

- Absorption of energy, at appropriate wavelengths, by the various gaseous molecular constituents of the atmosphere, principally H_2O, O_2, CO_2 and O_3 (ozone)
- Scattering by dust and aerosols* suspended in the atmosphere.

---

*An aerosol is a suspension of finely divided solid or liquid matter in a gaseous medium.

Molecules have many distinct internal energy states, corresponding to different levels of rotational, vibrational, or electronic energy. When a molecule is struck by electromagnetic radiation at a wavelength corresponding to the exact energy ($E = hc/\lambda$) required to raise the molecule to a higher energy state, that amount of electromagnetic radiation is absorbed. This energy is later reradiated at the same wavelength, as the molecule returns to the original state of lower energy. However, the energy is not necessarily radiated along the original path of propagation. Hence, the original radiation is said to have been absorbed. This process is responsible for the sharp absorption peaks seen in Fig. 3.4-1 (a) and (b).

Molecular motion and collisions cause broadening of the absorption lines, so that electromagnetic radiation at wavelengths near the critical wavelength is also absorbed, but to a lesser extent. In addition, closely spaced energy transitions give rise to closely spaced absorption lines and a broader, more continuous region of absorption. Both these situations apply, for example, to the absorption due to O_2 at 5 mm (60 GHz), and, on a broader scale, to the absorption continuum caused by H_2O absorption lines from 3 mm (100 GHz) down to approximately 10 μm. The latter part of the spectrum is not included in Fig. 3.4-1 because the atmosphere is essentially opaque in this part of the spectrum, except for a few narrow windows suitable only for very limited propagation. In addition, sensing technology is not well developed in the far-far infrared.

The atmosphere is completely opaque for wavelengths shorter than 0.25 μm -- that is, from the ultraviolet (UV) through x-rays and gamma rays. This prevents biologically harmful radiations, originating in cosmic rays and the sun, from reaching the earth's surface -- except for UV between

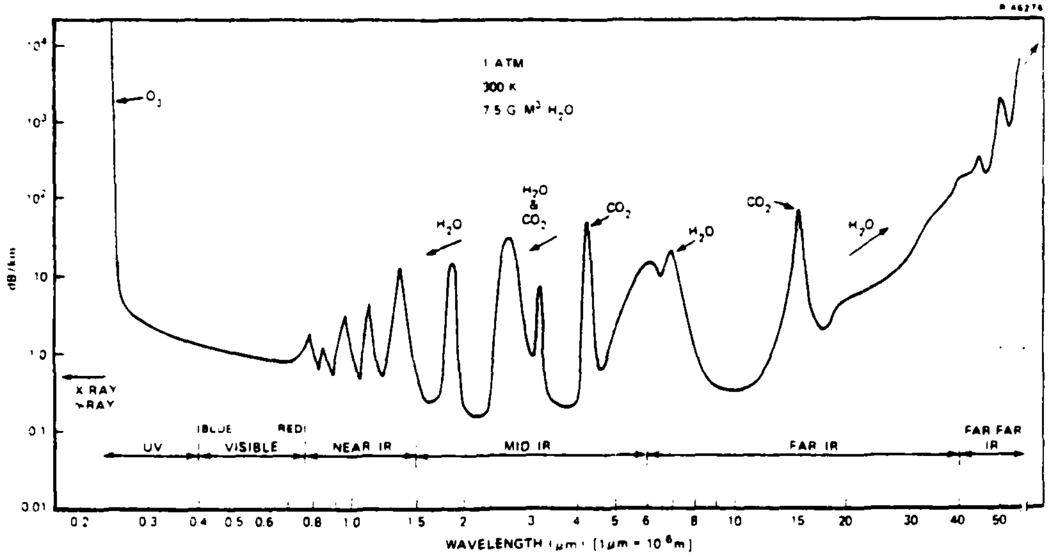


Figure 3.4-1a   Atmospheric Attenuation from UV to Far-Far IR

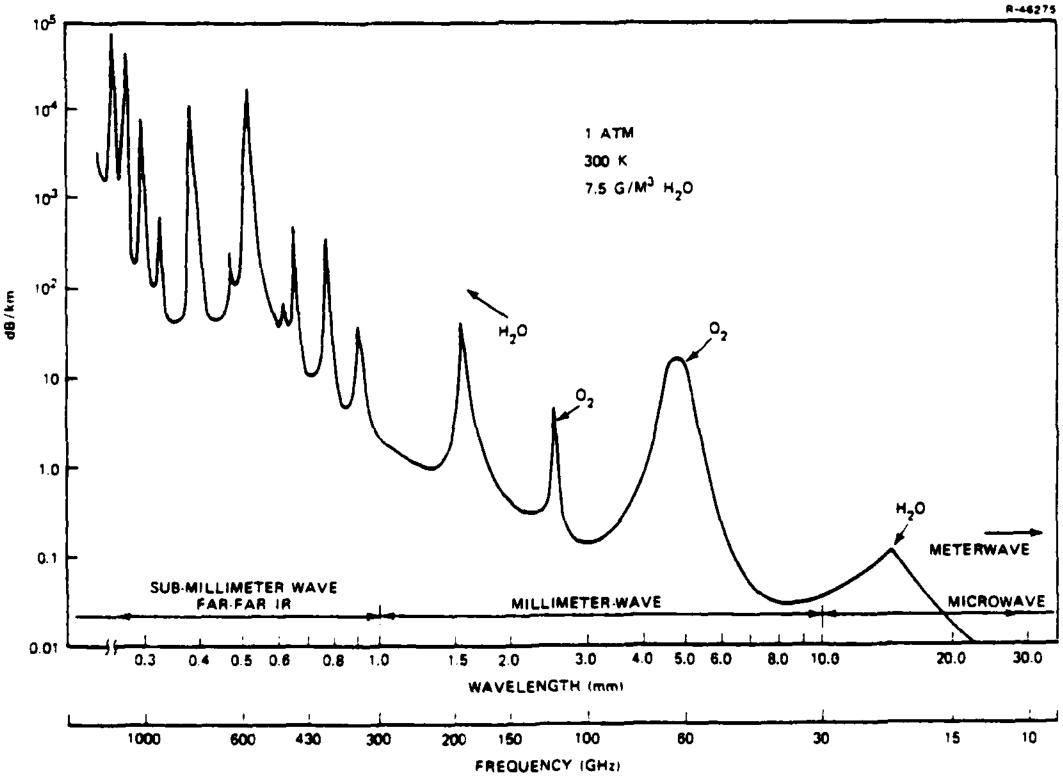


Figure 3.4-1b   Atmospheric Attenuation from Sub-Millimeter to Microwave

0.25 and 0.4 μm which can cause sunburn and skin cancer in humans. UV at wavelengths shorter than 0.25 μm is blocked principally by the ozone layer in the upper atmosphere.

In the visible-light window, blue light undergoes slightly more atmospheric attenuation than red light. This results from what is known as Rayleigh scattering. Tiny dust and liquid particles suspended in the atmosphere scatter electromagnetic radiation. In the case of Rayleigh scattering, where the wavelength is long compared to the dimensions of the suspended particles, the scattering (and therefore attenuation) properties of the particles vary inversely as the fourth power of the wavelength. In other words, as the wavelength decreases, the scattering and net attenuation increase rapidly. This explains why the sky is blue. The blue part of the sun's light is scattered randomly by particles in the atmosphere and reaches the ground from all directions; the sky everywhere other than in the direction of the sun then appears blue. Sunsets are red because of similar scattering phenomena. At low elevation angles, the sun's light has to traverse a much longer path through the atmosphere, with the result that only the deepest red light has not been scattered (and attenuated) away.

In addition to the very narrow windows seen at various parts of the spectrum in Figure 3.4-1, there are a number of broader windows commonly used in reconnaissance and remote sensing. These are

- Visible Light
- Mid-IR, from approximately 3 to 5 μm (1.0×10^{14} to 6.0×10^{13} Hz)
- Far-IR, from approximately 8 to 14 μm (3.75×10^{13} to 2.14×10^{13} Hz)

- Approximately 1.36 mm (220 GHz)
- Approximately 2.14 mm (140 GHz)
- Approximately 3.19 mm (94 GHz)
- Approximately 8.57 mm (35 GHz)
- Approximately 15 mm to 20 m (20 GHz to 15 MHz).

The low-frequency cutoff for the 15 mm to 20 m window is the characteristic plasma oscillation frequency of the ionosphere.

3.4.1 Attenuation Effects of Haze, Fog, and Adverse Weather

This section presents the effects on electromagnetic propagation of the larger atmospheric aerosol particles found in haze, fog, clouds, rain, and snow. The relative importance of absorption versus scattering is strongly dependent on the wavelength of the radiation compared to the typical diameter of the aerosol particles being considered.

As mentioned earlier, the case where the wavelength is long compared to the dimensions of the aerosol particle is known as Rayleigh scattering; here the scattering increases as the fourth power of the frequency (or the inverse fourth power of the wavelength). As the wavelength of the incident radiation decreases and becomes comparable to the size of the particle, the problem becomes more complex, with a maximum scattering efficiency being reached when the size of the particle is approximately one wavelength. This situation is called resonant scattering. Finally, when the wavelength becomes very short compared to the aerosol dimensions (the optical scattering region), the scattering and net attenuation asymptotically approach a limiting or geometrical value.

These effects are shown in Figure 3.4-2 for a variety of atmospheric aerosols across the wavelength region covered by Fig. 3.4-1. As in Fig. 3.4-1, these curves represent approximate values of attenuation and are intended to present overall trends instead of exact numerical values. Moreover, the attenuation values indicated by these curves must be added to the clear air attenuation given by Fig. 3.4-1 to obtain actual values as a function of wavelength for a particular atmospheric condition. Thus, even though submillimeter waves at approximately 200 μm appear ideal for propagation through thick fog (400 m visibility), the natural attenuating properties of the clear atmosphere at those wavelengths still render them unsuitable for all but very short distances.

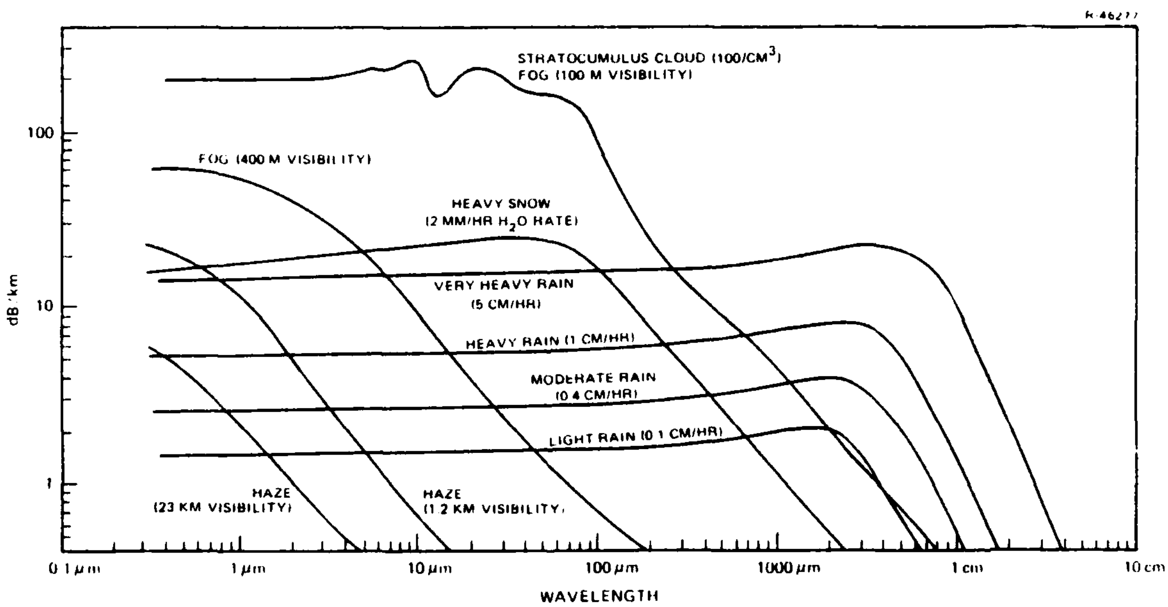


Figure 3.4-2 Atmospheric Attenuation from Haze, Fog, and Adverse Weather

Generally, the peak of an attenuation curve in Fig. 3.4-2 lies at a wavelength corresponding roughly to the size of the aerosol particles most commonly found in the particular precipitation or environmental situation. The raindrop sizes are directly related to rainfall rates -- the higher the rainfall rate, the larger the raindrops that may be expected. With larger water droplets, the attenuation peak is reached at longer wavelengths. Also, there is more attenuation at the peak and at the shorter wavelengths.

As an example of the importance of the curves of Fig. 3.4-2 in sensor system planning, observe that for remote sensing through moderate rain (0.4 cm/hour), wavelengths longer than 3 mm (frequencies less than 100 GHz) are optimal. As a second example, infrared wavelengths are preferred over visible wavelengths for remote sensing through haze and light fog.

3.4.2 Backscatter from Atmospheric Aerosols

Active imaging systems involve scanning a beam of radiation across the target area and sensing the return signal variations in intensity as a function of angular position and range to the target. The effects of atmospheric aerosols and clear-air attenuation are more complicated for such an imaging system.

For example, the clear-air attenuation curve in Fig. 3.4-1 shows the one-way path loss. The signal from an active system has to traverse a total distance equal to twice the range between sensor and target, so that the total path loss is greater than that for a passive system at the same range. Since radiation transmittance, τ, along a path is given by

$$\tau = e^{-\alpha r/4.34} \qquad (3.4\text{-}1)$$

where α is the attenuation in dB/km, and r is the path length, doubling the path length is equivalent to changing the transmittance from τ to τ^2. The transmittance lies between zero and one, so $\tau^2 < \tau$. A system can overcome this penalty by increasing its transmitted power to guarantee sufficient return power for detection by the sensor, as dictated by the requirement that an active system must be designed so that its signal strength, upon reflection from the target, is significantly greater than any natural thermal radiation the target is emitting. (Thermal radiation is discussed in Section 3.5.1.)

A potentially more serious problem for active systems is backscatter from atmospheric aerosols. The attenuation curves of Fig. 3.4-2 result from the scattering of the incident radiation by aerosols of various sizes. The part of the incident radiation scattered back into the field of view (and potentially masking the radiation returning from the target) is said to be backscattered. This effect is responsible, for example, for the difficulty encountered in driving an automobile at night in the fog. Most of the light from the headlights is backscattered by the fog particles, creating a widespread diffuse glow and reducing the visibility of the road beyond. As scattering decreases with increasing wavelength, fog lights emitting yellow light of longer wavelength can sometimes increase visibility.

At a fixed wavelength, the amount of backscattering is related to two parameters: the solid angle subtended by the transmitted beam of radiation as it enters the backscattering medium, and the total scattering volume (i.e., the solid angle of the beam multiplied by the path length through the backscattering medium). Reducing either the beam solid angle or the path length through a scattering volume reduces the amount of radiation backscattered.

When an active imaging system can provide range information, the range resolution (as determined by the duration of a transmitted pulse) along with the beam solid angle determines the scattering volume. By increasing the range resolution (shortening the pulse duration), the scattering volume per pulse is decreased and backscatter is reduced. In addition, the transmitted pulses may be encoded with a particular waveform or polarization that is not reflected by small aerosol particles as well as by the larger, flatter target surfaces. These techniques are commonly referred to as clutter suppression and can significantly improve an active system's operation in adverse weather.

A continuous wave (CW) system (like the illumination from car headlights) is much more severely affected by backscatter, in part because the scattering volume is much larger. Active imaging CW systems thus generally cannot be used during adverse weather if the radiation is likely to be backscattered.

3.4.3 Scintillation and Atmospheric Turbulence

In any kind of imaging system, more sensitivity can be obtained by increasing the effective aperture of the receiver. An active system can increase its sensitivity by increasing its transmitted power; this may not be desirable, however, in covert applications of the imaging system where the system's transmitted radiation serves not only to illuminate the target area but also to announce the presence of the active system to other (possibly unfriendly) sensors in the target area. Very often, however, vehicle constraints limit the power and aperture size of an imaging system.

The aperture size of an imaging system also determines to a large extent the angular resolution achievable in the

image. The larger the aperture, the higher the angular resolution. According to the theory of diffraction, the angular resolution achievable (in radians) is approximately λ/D, where λ is the wavelength of the radiation and D is the aperture of the imaging system.

At infrared, visible light, and ultraviolet wavelengths, there is a limit to the amount of detail that can be seen on the earth's surface from great distances (for example, from orbit). Rays of light (IR, visible, or UV radiation) emanating from the target area are bent slightly by the earth's atmosphere as they travel to the imaging system. When a ray of light travels through the earth's atmosphere at an oblique angle, it bends or refracts because of the slightly slower speed of travel in the atmosphere, just as a ray of light is bent when it enters or leaves a piece of glass (see Fig. 3.4-3).

R-47184

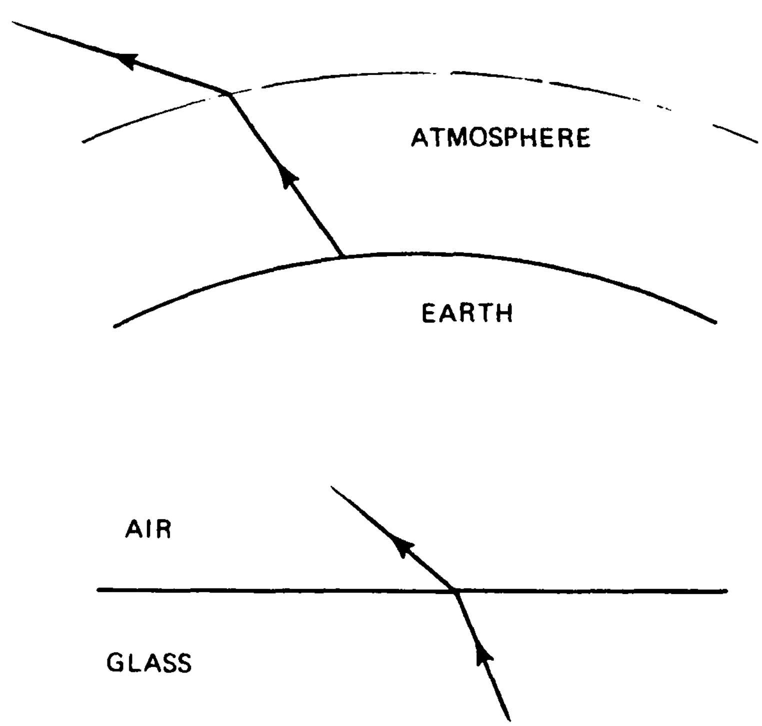


Figure 3.4-3 Refraction Example

Atmospheric refraction would have no effect on the clarity of the image formed by the sensor if the air were perfectly quiet; it would only shift the apparent position of the image by a small amount. The atmosphere is, however, continually in motion and, as a result, the position of the image keeps changing. This effect produces, for example, the twinkling of starlight, as well as the shimmering, distortions, and linear displacements observed when looking through a layer of water.

The blurring caused by atmospheric effects can be reduced by using a shorter exposure time (the time required for the sensor to produce the image), but it is necessary to compensate by increasing either the sensitivity or the aperture of the detection system. Other techniques for reducing the effects of atmospheric blurring include <u>adaptive optics</u> (the use of multielement apertures) and <u>speckle interferometry</u> (the use of signal processing methods to reconstruct the original flat wavefronts from the distorted wavefronts entering the sensing system's aperture).

Finally, it is emphasized that the atmospheric blurring of images is a separate phenomenon from atmospheric transparency. Atmospheric blurring is an indication of the steadiness of the earth's atmosphere which is called, technically, the <u>seeing</u>. <u>Transparency</u> is a measure of how clear, or transparent to electromagnetic radiation, the atmosphere is. It is quite possible for the transparency to be good but the seeing to be very poor, or for the transparency to be bad (a hazy sky in visible light, for example) but seeing excellent.

CHAPTER FIVE

INFRARED IMAGING SYSTEMS

3.5.1 Infrared and Thermal Radiation

Infrared (IR) designates electromagnetic radiation lying at wavelengths longer than the visible spectrum (> 0.75 μm) and shorter than microwaves/millimeter-waves (< 1 mm). IR is classified according to its wavelength as near (0.75 to 1.5 μm), middle (1.5 to 6.0 μm), far (6.0 to 40 μm), and far-far or sub-millimeter (40 μm to 1 mm).

Any object is a source of thermal radiation, which is a result of the thermal motions of the atoms and electrons of which the object is composed. Furthermore, the spectrum of this radiation is continuous, because it results from the motions of many atoms, each of which interacts with its nearest neighbors. In objects heated to 1500 K and above, the atoms have sufficient energy to emit radiation at the visible wavelengths to which the human eye responds. However, an object at a lower temperature, even room temperature (~300 K), still emits thermal radiation, though it is not visible. Such radiation can be detected by infrared techniques. As the temperature of an object is lowered, the quantity of radiation decreases and the wavelength of peak emission shifts to longer wavelengths (lower frequencies).

Figure 3.5-1 is a plot of the infrared energy distribution of a perfect radiator (blackbody*) at various temperatures,

*The term blackbody is used to describe a hypothetical object that absorbs all radiation incident upon it.

described quantitatively by Eq. 3.5-1, known as Planck's radiation law.

$$w_\lambda = 4\pi c^2 h\lambda^{-5}(e^{hc/\lambda kT} - 1)^{-1} \quad (3.5\text{-}1)$$

where

c = speed of light (3×10^8 m/sec)

h = Planck's constant (6.6×10^{-34} joule sec)

k = Boltzmann's constant (1.4×10^{-23} joule/K)

λ = wavelength (m)

T = temperature of the object (K)

Figure 3.5-1 plots Eq. (3.5-1) for various values of T.

A true blackbody, an object that emits the maximum possible amount of radiation for its temperature and absorbs all energy incident upon it, does not exist in the real world. In interacting with physical objects, electromagnetic radiation may be reflected, absorbed, or transmitted. The reflectance, absorptance, and transmittance of a substance are related by Eq. (3.5-2):

$$\rho + \varepsilon + \tau = 1 \quad (3.5\text{-}2)$$

where ρ is reflectance, ε is emissivity (equivalent to absorptance), and τ is transmittance. A perfect blackbody, for example, would have $\varepsilon = 1$, and $\rho = \tau = 0$. A perfect vacuum free of EM fields would have $\tau = 1$, and $\rho = \varepsilon = 0$. All imaging infrared (IIR) systems, active or passive, take advantage of certain of these characteristics and are adversely affected by others.

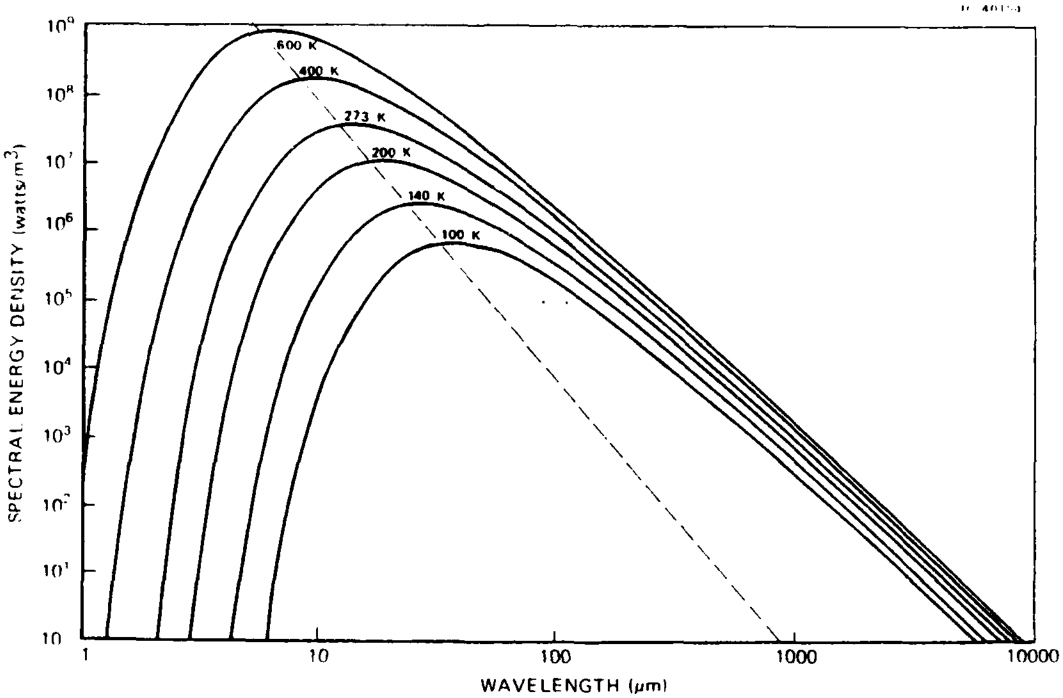


Figure 3.5-1 Radiated Energy vs Temperature

Passive detection systems can distinguish among objects of the same temperature because each object has a different emissivity and accordingly emits different amounts of radiation. Radiation at the same wavelengths but at various intensities from surrounding materials (terrain, sky, etc.) is a source of interference since such radiation can be confused with the radiation from the object. An active system, on the other hand, relies on the reflectivity of the target to produce a detectable return signal. With an active system, objects with high reflectivities are detected most easily.

3.5.2 Infrared Sensors

Imaging infrared sensors have been under military development since about 1950. They are variously referred to

as thermal imagers, thermal viewers, IR line scanners, or IR focal-plane mosaics. Often, trade names or military system acronyms like FLIR, for Forward-Looking Infrared, are used. An IR sensor consists of an IR optical lens assembly (transparent to the wavelengths of interest), IR detector, electronics, and a recording device and/or display. An example of the widely used Common Module FLIR system* is illustrated in Fig. 3.5-2.

Most detectors fall into one of two general classes: thermal detectors and photon detectors. The earliest infrared detectors were thermal detectors, in which the energy of the absorbed radiation raises the temperature of the detecting element. This causes measurable changes in temperature-dependent properties of the detector. For example, a bolometer detector changes its resistance in response to incident radiant energy; a pyroelectric detector changes its capacitance in response to thermal energy. Thermal detectors are well suited for broadband detection, can operate at room temperatures, but have relatively slow response times (and, consequently, cannot be used in rapid-scan imaging techniques). In addition, two-dimensional thermal imaging systems often suffer from poor spatial resolution.

In photon detectors, the incident radiation excites electronic transitions which change the quantum state of electrons in the detector's sensitive element; such detectors are sensitive to the number of incident photons. These detectors have a more limited spectral sensitivity range than thermal detectors, but are characterized by relatively fast response times and high sensitivities. The photon detectors, using

*The Common Module FLIR, used by the Army, Navy and Air Force, consists of interchangable modules which are adaptable to a variety of military missions.

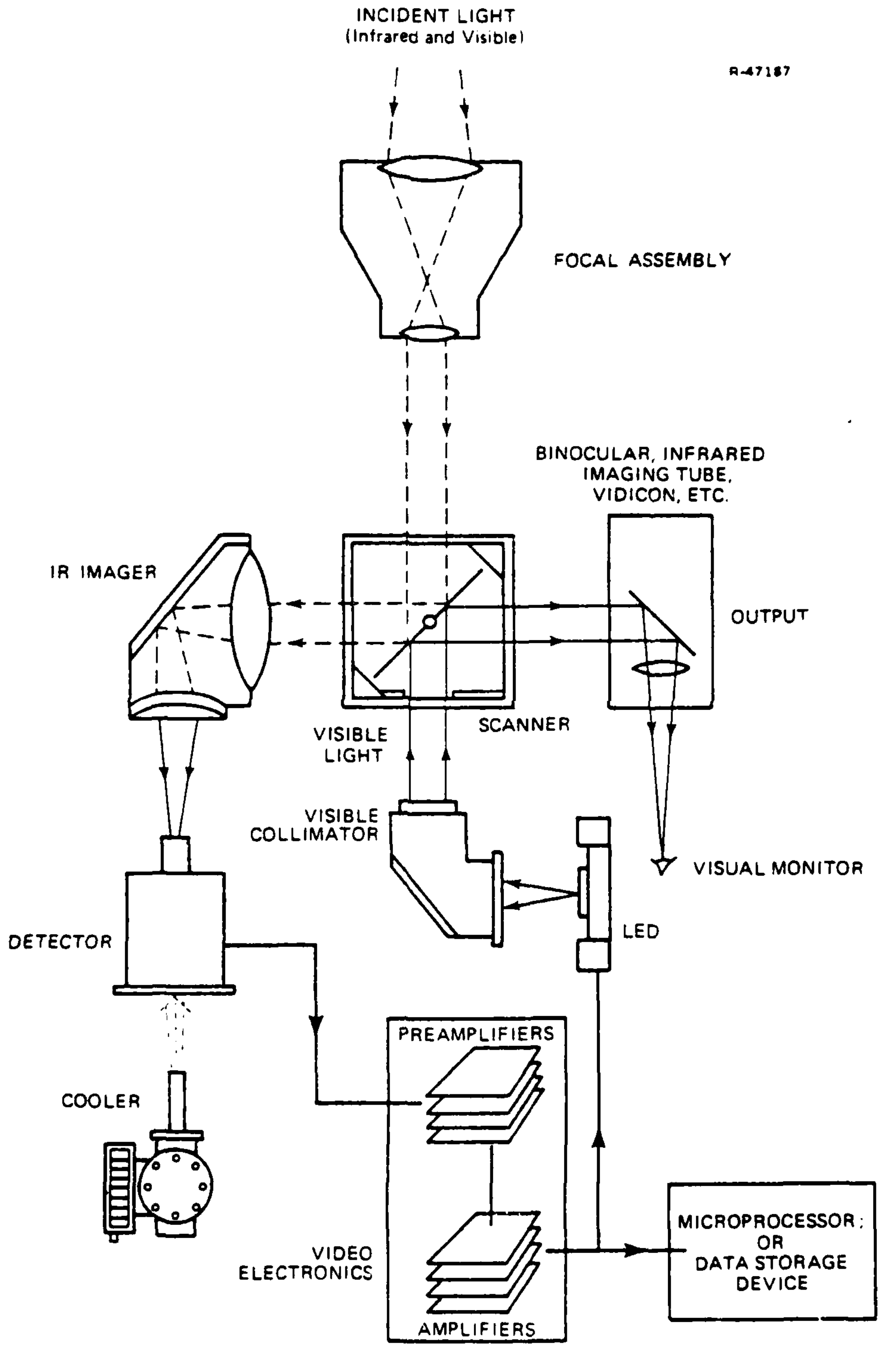


Figure 3.5-2 Common Module FLIR System

solid-state crystals such as indium antimonide, are physically more rugged than thermal detectors, and have therefore been more widely used in military reconnaissance projects. However, the solid-state devices generally require cryogenic cooling to obtain adequate noise-free sensitivity.

Detector materials have various spectral responses and are selected in accordance with favorable atmospheric transmission windows and the radiation characteristics of the area or targets being imaged. Extremely hot targets, for example, are best sensed by detectors sensitive in the 1.0 to 6.0 μm band, while targets at near-ambient temperature are best sensed with detectors sensitive to longer wavelengths. Figure 3.5-3 shows the spectral responses of commonly used detector materials. The acronym BLIP in Fig. 3.5-3 refers to Background Limited Performance with a 295 K background temperature over a

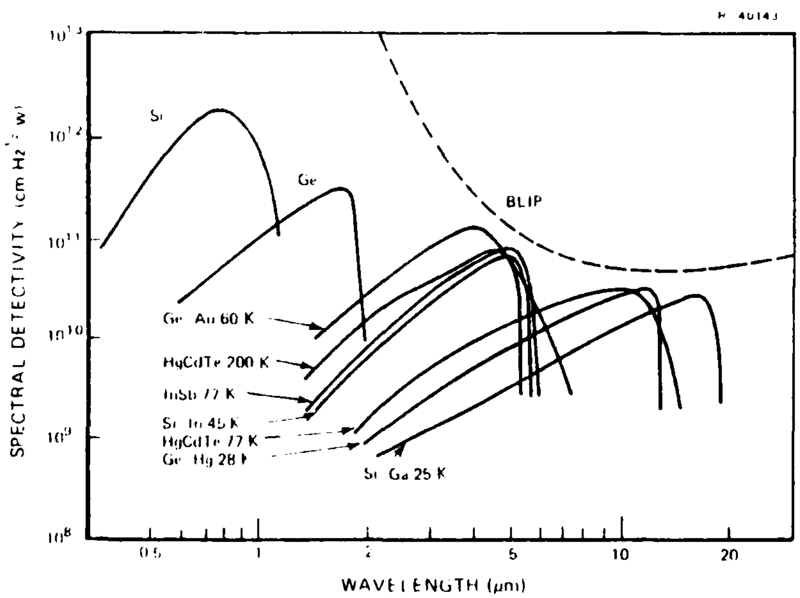


Figure 3.5-3 Spectral Sensitivity of Detector Materials

hemispherical field of view;* an ideal photon detector operating in such a temperature environment could, because of the background noise, never exceed the performance indicated.

In addition to the sensitive solid-state photon detectors, various IR-sensitive vidicon-type tubes and image intensifiers are also used. These devices operate generally in much the same way as a conventional television tube, except that the detector material is sensitized to IR wavelengths. A commonly used detector is a silicon diode array (approximately 10 by 12 mm), formed from a single crystal of silicon and scanned with a conventional electron beam. This device is useful principally in the near-IR spectral region.

Another type of IR vidicon is the <u>pyroelectric</u> vidicon, which is like a conventional vidicon, except that the infrared-transmitting faceplate is usually a germanium disk, transparent to radiation beyond about 2 μm, with an antireflection coating for maximum transmission in the 8 to 14 μm spectral band. The radiation-sensing retina is a single crystal disk of an insulating pyroelectric material like triglycine sulphate (TGS) with a semitransparent electrically conducting signal electrode formed by evaporation on the surface next to the faceplate. For a 25 mm tube, the active image area is 16 to 18 mm in diameter, and the retina is approximately 30 μm thick. To produce a TV-compatible thermal image of a stationary scene with a pyroelectric vidicon, the radiation pattern must be modulated by a chopper wheel (in synchronization with the electron beam scanning pattern), because the pyroelectric material senses only changes in temperature -- <u>not</u> changes in photon flux. Thus, in alternate frames, the retina views first the scene image, then the chopper blade. If the scene area is hotter

*Note that operational sensors usually have narrower fields of view.

than the chopper, the retina warms during the scene exposure. producing a corresponding positive charge pattern. It cools during a chopper exposure, producing a negative charge pattern. A low-energy scanning electron beam reads the charge pattern from the retina surface to produce a conventional TV-like image.

A major disadvantage of IR-sensitive vidicon tubes is their comparatively slow response time. In addition, their infrared sensitivity is usually coarser than that provided by such detector materials as HgCdTe. The vidicons, however, because of their electron-beam scanning feature, are generally able to provide imagery with excellent angular resolution. Since the current emphasis in the development of military systems is toward very sensitive imagery of rapidly changing or moving scenes, another type of detector, the Charge Coupled Device (CCD), is receiving much attention.

3.5.3 Passive Infrared Imagery

Most passive infrared imaging systems fall into two main categories:

- Scanning
- Staring.

In a scanning system, the detector (or, more commonly, an array of detectors) is swept across the field of view. Alternatively, the field of view can be scanned across the linear array. In a staring system, the entire field of view is filled with a mosaic of detectors, and scanning techniques are usually not required. Detector sensitivity, S, is given by the relation

$$S = \frac{K\, N_{det}\, A_{det}}{A_{fov}} \qquad (3.5\text{-}3)$$

where

A_{det} = area subtended by the detector

A_{fov} = area corresponding to desired field of view

N_{det} = number of detector elements

K = a proportionality constant

The sensitivity of a scanning system is proportional to the length of time of the scan. This is generally less than that of a staring sensor (where each detector element can view its portion of the scene continuously). However, the technique of time delay and integration (TDI) can be used to increase sensitivity for scanning systems; the shift rate for the output signal delay line is synchronized with the scan rate. Figure 3.5-4 illustrates the concept of TDI imaging.

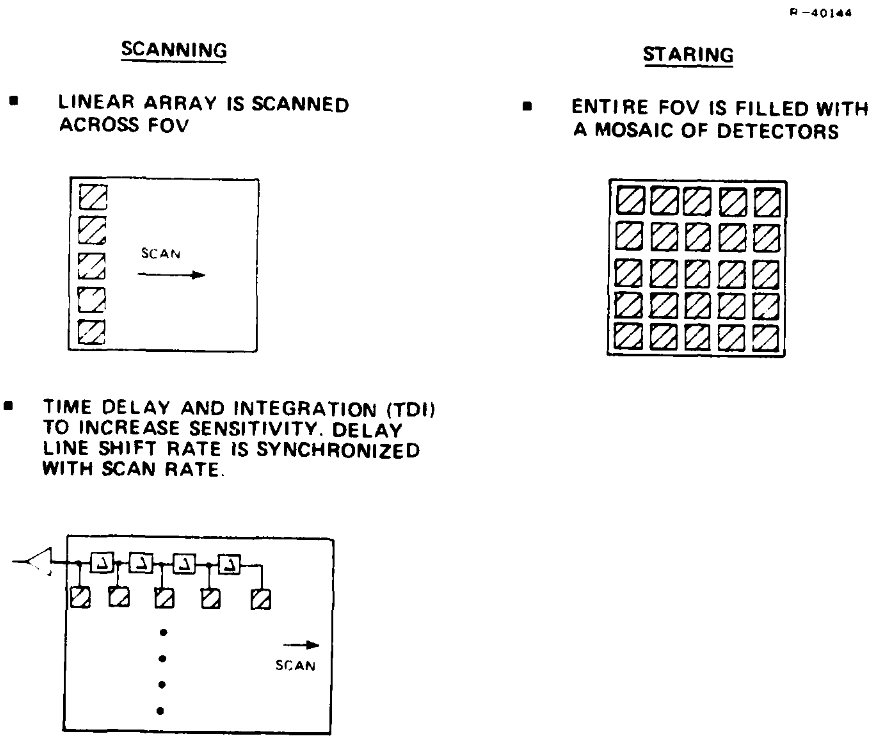


Figure 3.5-4 TDI Scanning Principle

Staring-array systems have a number of advantages over scanning-array systems. In addition to being more sensitive, they are potentially cheaper (in part because they usually require no moving parts), and are less sensitive to the effects of bright spots in the field of view. However, large IR mosaic arrays (more than 64 by 64 elements) are not currently available, principally because the array size is limited by the mechanical difficulties involved in bringing many wire leads (one for each element) out of the sensor. However, advances in solid-state technology are making it possible to fabricate detectors and readout electronics together, eliminating the need for wire leads in the focal plane and allowing the development of very large arrays. Such arrays will have obvious advantages as sensors on reconnaissance aircraft and satellites.

Because two-dimensional array technology is still in the early stages of development, most IIR systems used at the present time for navigation or reconnaissance are of the scanning type. Imaging methods using only one detector are shown in Figure 3.5-5. If the sensor is mounted on an aircraft or missile, imagery of the ground beneath (downlooking imagery) can be obtained with the sensor scanning in a direction perpendicular to the direction of flight (Figure 3.5-5a), with the forward motion of the vehicle moving the scan downtrack. If the sensor platform is stationary, or looking forward along the vehicle flight path, the sensor itself must scan a two-dimensional raster scan (Fig. 3.5-5b).

Imaging methods using a linear array are shown in Fig. 3.5-6. Downlooking imagery can be obtained in a push-broom mode (Fig. 3.5-6a), involving no moving parts in the sensor system. A linear array is oriented crosstrack, and the forward motion of the vehicle carries the instantaneous image

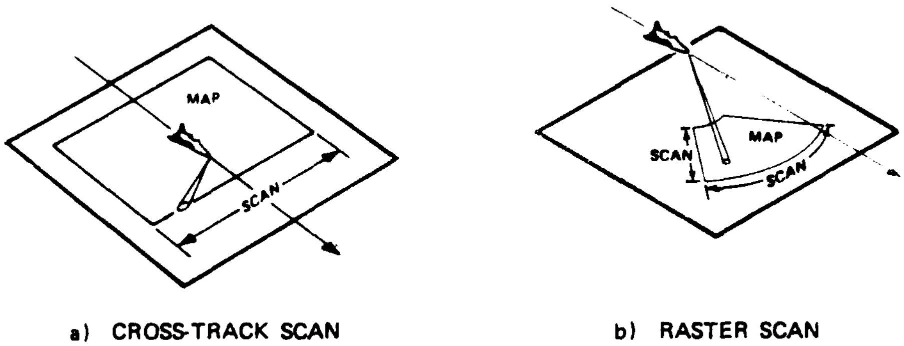


Figure 3.5-5 Single Sensor Scanning Techniques

line downtrack. Forward-looking imagery can be obtained by orienting the array vertically and scanning in azimuth (Fig. 3.5-6b), or by orienting the array horizontally and scanning in elevation (Fig. 3.5-6c). Either of these methods can produce a TV-compatible image. In fact, the FLIR system included as part of the cockpit instrumentation for a number of aircraft produces such an image by the method shown in Fig. 3.5-6b.

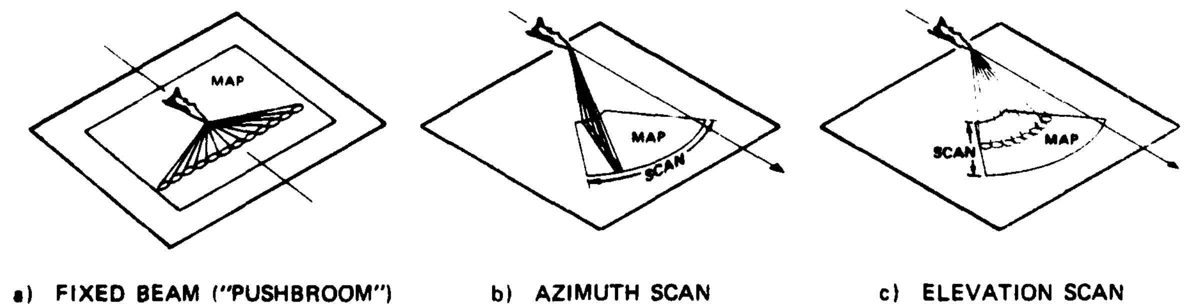


Figure 3.5-6 Array Scanning Techniques

3.5.4 Active Infrared Imagery

Active infrared (laser radar) systems use laser radiation to illuminate the target area. These systems can provide range information as well as two-dimensional imagery. Common IR laser sources are listed in Table 3.5-1. In scanning the target area, the laser and the receiving detector (which can be a single element, a linear array, or a small mosaic array) may move together, or the receiver may be fixed in one or both dimensions.

In pencil-beam laser systems, the scanning techniques are the same as those shown in Fig. 3.5-5a,b. If the laser system is airborne, a lens or mirror system elongates the beam in the vertical dimension, producing a so-called fan beam; this corresponds to Fig. 3.5-6b, where the downrange divisions in the beam would be implemented using range gates.* In the fan-beam system, the two dimensions of the imagery are azimuth and range.

The two most important characteristics of detectors for active infrared imagery are sensitivity and speed. The commonly used detectors are the photomultiplier, silicon PIN-junction photodiode, silicon and germanium avalanche photodiode, photoconductor, and photovoltaic detector.

The photomultiplier resembles most closely the ideal photodetector. Photons incident on a photosensitive surface produce photoelectrons, which are accelerated into dynode targets to produce secondary electrons. Several stages produce

*In a range gate system the return time of the laser signal is measured to provide the range to the observed object.

TABLE 3.5-1
COMMON IR LASER SOURCES

LASER	TYPE	WAVELENGTH (μm)	MODE OF OPERATION
HeCd	Gas (Ion)	0.325-0.44	CW
Argon	Gas (Ion)	0.35-0.53	CW, Pulse
HeNe	Gas (Neutral Atom)	0.63, 1.15, 3.39	CW
Ruby	Crystal	0.69	Pulse
GaAs	Semi-Conductor	0.82	CW
GaAs	Semi-Conductor	0.9	Pulse
Nd:YAG	Crystal	1.06	CW, Pulse
Nd:YAG×2	Crystal	0.53	CW, Pulse
HF	Gas	2.8-5.0	Pulse
DF	Gas	3.8	Pulse
CO	Gas	5.1-5.3	Pulse
CO_2	Gas	10.6	CW, Pulse

significant current gain (to suppress thermal noise of the following amplifiers) in a basically noiseless process. There are many photosurfaces available, but all are limited to the visible and near IR. Photomultipliers require high voltage supplies, but are rugged, simple, and suitable for field use. Under most daylight conditions, the systems are background limited. Under ideal conditions, single photoelectron detection is possible.

3.5.5 IIR Scanning Considerations

There are a number of scanning techniques commonly used in IIR systems. These fall into two main classes:

- Those that optically scan the field of view across a fixed detector
- Those that move the detector across the focal plane.

Each of these systems may be mounted on multiaxis gimbals to allow coarse positioning of the desired image frames; however, the actual scanning required to produce the image is rarely done this way.

Some examples of basic scanning techniques are shown in Fig. 3.5-7. Figure 3.5-8 illustrates a carrousel optical scan system designed by Hughes Aircraft Company. It uses the technique shown in Fig. 3.5-7a.

Regardless of the actual technique used, a scanning IIR system produces two-dimensional radiation intensity information by sampling the target area in discrete lines. When the target is covered by the width of a single scan line (corresponding to the angular area viewed by a single detector element), it is imaged as a dot the size of the scan line. However, if the target is larger, more scan lines or detector elements cross it. As the number of scan lines or pixels crossing it increases, the object is more easily recognized.

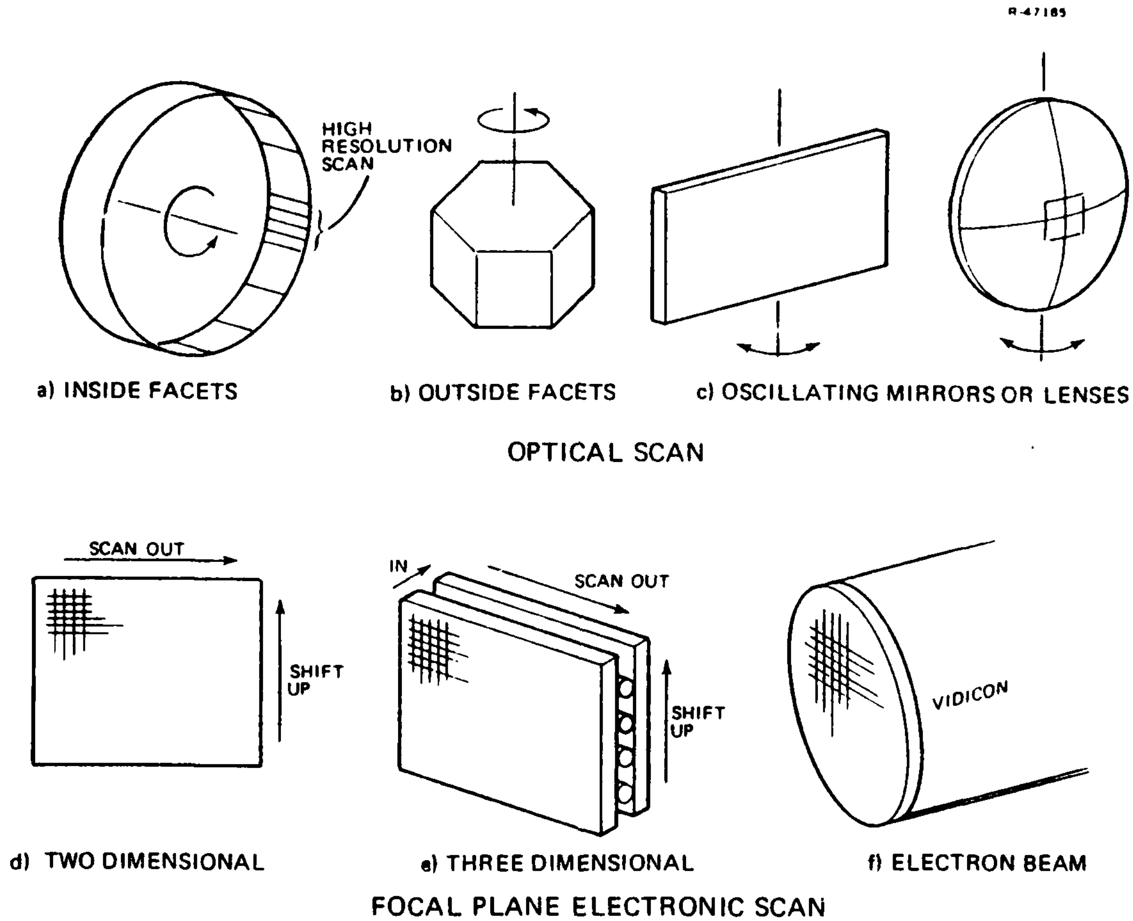


Figure 3.5-7 Basic Scanning Techniques

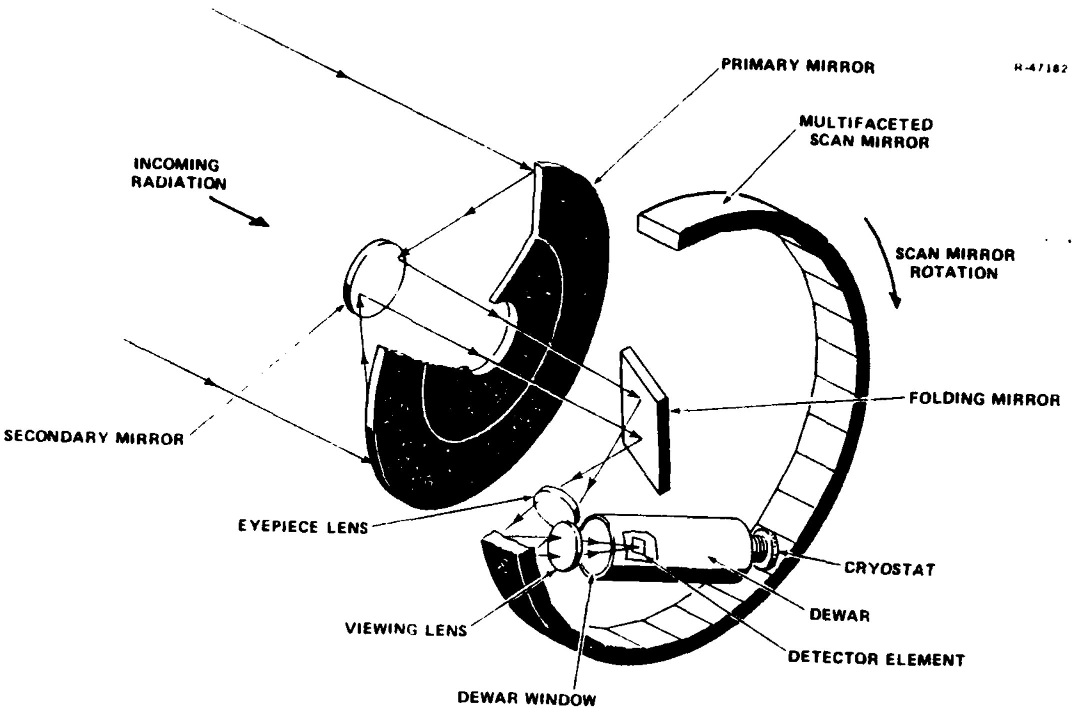


Figure 3.5-8 Carrousel Optical Scan

CHAPTER SIX
MICROWAVE AND MILLIMETER-WAVE IMAGING SYSTEMS

3.6.1 Microwave and Millimeter-Wave Radiation

Microwave (MW) and millimeter-wave (MMW) designate those wavelengths of electromagnetic radiation longer than far-far infrared or submillimeter (longer than one mm) but shorter than one meter.*

Thermal radiation, as described in Chapter Five, is emitted at millimeter, microwave, and infrared frequencies (see Fig. 3.5-1). There is, however, a wide disparity in the emissivities and reflectivities of objects at MW and MMW frequencies. Soil, for example, may have a very high emissivity and a very low reflectivity. Metal, on the other hand, has virtually no emissivity and therefore a unity reflectivity at these frequencies. In other words, different materials at the same temperature emit significantly different amounts of thermal radiation at MW and MMW frequencies. For objects at room temperature, however, the emission of thermal radiation is extremely low, particularly when compared with the IR emission at moderately elevated temperatures. In the past, this has limited the use of passive MW or MMW radiometry to radio astronomy, specialized remote sensing satellites, and research laboratory environments. However, receiver technology and MMW techniques have improved dramatically since the late 1960s. Sensitive and compact receivers have been developed for use in passive and

---

*The corresponding frequency range is 300 GHz to 300 MHz.

active systems. Passive radiometric imagery from aircraft or satellites is now used routinely for reconnaissance mapping and remote sensing. MW and MMW sensing is especially advantageous for situations where covert information is sought, where there is a wide variation in the expected emissivity/reflectivity values of the surface to be imaged, and where adverse weather capabilities are desired (principally at microwave frequencies; see Fig. 3.4-2). The short wavelengths of the millimeter band, while more vulnerable to rain and cloud attenuation, make possible small and light electronic hardware, as well as relatively good image-resolving capabilities.

There are several frequency ranges where clear-sky attenuation of MW and MMW radiation is minimized. These atmospheric windows are shown in Fig. 3.4-1b, and are centered roughly on the following frequencies: 35 GHz, 94 GHz, 140 GHz, 220 GHz. Moderate to strong attenuation occurs at 22.235 GHz, 183.3 GHz, and 323.8 GHz for water vapor; and at 118.8 GHz, as well as between 50 and 70 GHz, for oxygen.

Active mapping/imaging systems designed to operate over short ranges and in covert situations can take advantage of the attenuation provided by the MMW atmospheric absorption lines. For example, a low-altitude or short-range radar mapping sensor could operate close to 60 GHz, the peak of the oxygen absorption region. Atmospheric attenuation could be used to prevent the transmitted signal from propagating farther than desired through antenna sidelobe emission. Such a radar system would be virtually undetectable by any ground-based sensor unless it happened to illuminate the enemy's sensor directly.

3.6.2 MW and MMW Sensors (Passive)

The most significant difference between MW and MMW imaging systems, on the one hand, and optical systems (i.e., UV, visible-light, and IR) on the other, is that they use an antenna and a radio receiver as a sensing device, instead of a photon detector element or photographic film. The solid-state detector devices currently being produced (such as Schottky barrier diodes) convert the comparatively high-frequency incoming MW or MMW radiation, focused by the collecting aperture, to lower intermediate frequencies, usually less than 2 GHz, which can be manipulated easily in radio frequency circuits.

Imaging with a MW or MMW system is analogous to an optical system with only one detector element at the focus. In order to produce an image of a target area, the field of view illuminating the detector element must be moved -- either by scanning the entire optical system across the target area or by moving the detector element over the focal plane (if the instantaneous total field of view provided by the optics is sufficiently large). As mentioned in Section 3.4.3, the angular resolution obtainable with any imaging system is λ/D, where λ is the wavelength and D is a characteristic dimension of the imaging system. At millimeter wavelengths, this value is at least 100 times smaller than that obtainable with a 10 μm system of the same aperture. On the other hand, range measurement, if obtained with an active system, can have as high a resolution as the pulse-transmitting and signal-processing electronics allow; it is not related to aperture diameter.

The development of very sensitive MMW receivers has heightened the interest of military system planners in MMW radiometric reconnaissance, mapping, navigation, and tracking systems. One advantage of a passive radiometer over an active

radar system is that it is essentially impossible to detect. Another is that the range dependence is inversely proportional to target range squared, comparing favorably with the inverse fourth power dependence of radar. Also, there is no explicit signal interference problem since the radiometric signal is unstructured. The disadvantage is that even though radiometers consume less power and occupy less space than an equivalent active detection system, they yield less target information.

3.6.3 Passive MW and MMW Imagery

High-altitude aircraft or satellites can carry MW or MMW radiometers to map features on the earth's surface. Specific examples are discussed in Unit Four. Figure 3.6-1 illustrates the various sources contributing energy to the antenna of a radiometer looking at the earth. These include

- Emission by objects and features on the earth's surface
- Direct emission by the atmosphere
- Reflection of atmospheric emission from the earth's surface.

Also shown is the effect of atmospheric absorption in the transmission path from ground to antenna.

Emissivity values for natural and artificial objects at MW and MMW wavelengths range from almost one (e.g., earth) to nearly zero (for most metals). An object's emissivity is often dependent on a number of additional factors as well. For example, changing the water content of many materials (such as soil, sand, and snow) can significantly change their emissivities. The emissivity of a material may be much larger for vertically polarized radiation (radiation polarized normal

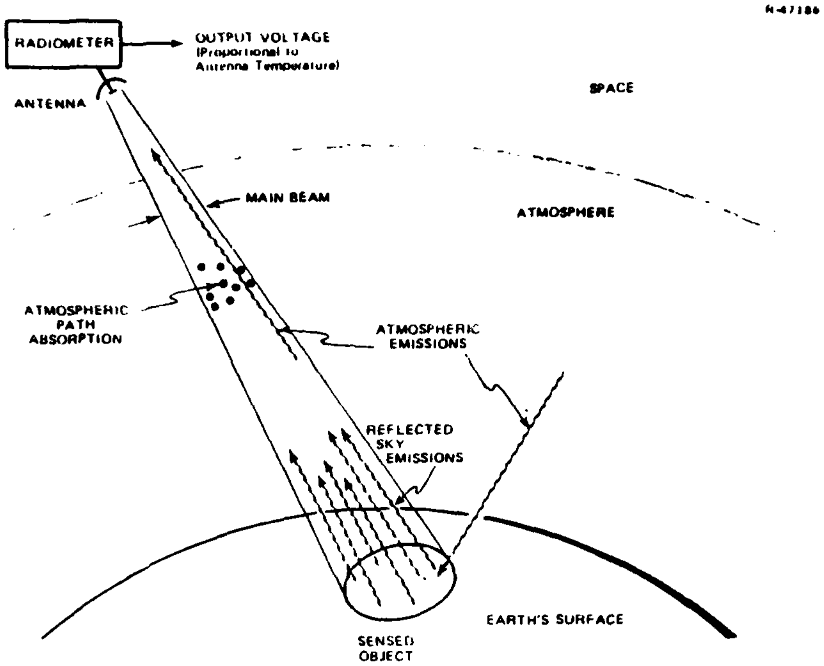


Figure 3.6-1 High Altitude Radiometer Measurement

to the plane of the material's surface) than for horizontally polarized radiation (radiation polarized parallel to the plane of the material's surface). These factors can cause difficulty in interpreting MW and MMW imagery obtained from either active or passive systems, arising from the wide and often unpredictable variations in emissivity and reflectivity values. Because of these problems inherent in passive MW and MMW imagery, many military radiometric reconnaissance, mapping, and tracking systems measure contrast temperature, discussed below, instead of actual radiometric temperature.

Contrast Temperature[†] - Contrast temperature, ΔT, is defined as the difference between the antenna temperature

(†)This section contains material at a more advanced level than the rest of the text.

of a scene with a particular target present, and the temperature when that target is absent. In general, a radiometric detector indicates an antenna temperature T_T from an altitude, h, at a nadir or zenith angle θ to the target scene expressed by

$$T_T = (1 - \frac{1}{L})T_a + \frac{\varepsilon T_t}{L} + \frac{(1-\varepsilon)T_R}{L} \qquad (3.6\text{-}1)$$

where

$L = \exp\left(\int_0^{h \sec\theta} \alpha(z)\,dz\right)$, atmospheric path loss

T_a = average atmospheric temperature between the target and radiometer

T_t = target scene temperature

ε = emissivity of the target scene

$\alpha(z)$ = weighting function

T_R = reflected sky temperature

The term T_R is the antenna temperature resulting from sky emissions reflected by the target scene. It is a function of the angle of incidence and the emissivity of the target surface filling the beam. For a metal target surface filling the antenna beam, the term $\frac{\varepsilon T_t}{L}$ in Eq. (3.6-1) vanishes ($\varepsilon \cong 0$) and $T_R = T_{SKY}$, causing T_t to be much smaller than typical terrain background values.

If a target surface did not completely fill the antenna beam, then the contrast temperature would be reduced by an antenna <u>fill factor</u>, F, defined as

$$F = \frac{\Omega_T}{\Omega_A} \tag{3.6-2}$$

where

Ω_T = solid angle subtended by the target at the radiometer antenna

Ω_A = the solid angle of the entire antenna beam.

The contrast temperature measured by the radiometer can be derived from Eq. (3.6-1) by assuming a mapping scenario -- for example, a metallic surface surrounded by terrain. The terrain background alone would yield an antenna temperature of

$$T_B = (1 - \frac{1}{L})\ T_a + \frac{\varepsilon T_t}{L} \tag{3.6-3}$$

When the metal surface enters the antenna beam, it partially occludes the background, and contributes to the antenna temperature by reflecting energy from the sky. The radiometer then measures an antenna temperature:

$$T_t = F[(1 - \frac{1}{L})T_a + \frac{T_{SKY}}{L}] + (1-F)T_B \tag{3.6-4}$$

where the parameter T_R (from Eq. 3.6-1) equals T_{SKY}. The contrast temperature ΔT is

$$\Delta T = \frac{F}{L}\ (T_t - T_{SKY}) \tag{3.6-5}$$

In clear weather, this contrast temperature between two disparate scenes is large enough to be easily detected by passive MW or MMW systems.

3.6.4 MW and MMW Radar Sensors

The receiver characteristics of active (radar) MW and MMW imaging systems are similar to those of passive sensors. The transmitting functions of the system usually are provided by a radiation source illuminating the same antenna* used by the receiver electronics.

The cross-sectional shape of the transmitted beam of radiation is usually circular, as is provided by a standard paraboloidal reflector antenna or horn lens. However, the shape of the reflector antenna is sometimes designed to produce an elongated beam in the elevation direction, or to transmit more power in the upper part of the beam. In observations at any depression angle** other than 90 deg (corresponding to aiming straight down toward the ground) the upper part of the radar beam illuminates the ground at a greater range than the lower part of the beam. It is useful to direct more power to the upper part of the radar beam in order to keep the returned power the same at all ranges. Range values, if desired, are determined through range-gating techniques, usually involving pulsed rather than CW transmitters.

A growing variety of devices are used to generate the transmitted MW or MMW radiation. Although vacuum-tube power sources, such as traveling wave tubes or klystrons, are still employed, particularly at MW and lower frequencies, solid state devices such as Impatt Oscillators and EIOs (Extended Interaction Oscillators) are in common use at the higher MMW frequencies.

---

*A radar system at any wavelength with a separately located transmitter and receiver is often referred to as a bistatic or semi-active system. Such systems are used much more for tracking and guidance than for mapping or reconnaissance.

**The angle below the local horizon at which an airborne imaging sensor is aimed.

The solid state generators are especially important when flight-rated hardware and low power levels are required. An additional class of devices using relativistic effects to generate very high levels of MMW power is under intensive investigation. The principal applications may be for ground-based radar systems.

3.6.5 Active MW and MMW Imagery

Active MW/MMW imagery is similar to active infrared imagery in that range information can often be obtained along with the two-dimensional imagery. Active imagery, or radar, relies principally on the reflectivity of materials. The radar cross section (σ) is a figure of merit commonly used to describe the microwave reflection efficiency of a target of interest. It is based on the natural reflectivity of the surface. Also important is the percentage of incident radiation reflected back along the incident path relative to the amount reflected in other directions. Since an isotropic sphere scatters an oncoming beam uniformly in all directions, its cross section is given by

$$\sigma = 4\pi J/H \qquad (3.6.6)$$

where J is the radiation intensity from the target and H is the irradiance at the target. In contrast with the uniform sphere, cross section values for large complex structures vary rapidly with aspect angle, particularly at short wavelengths.

A major source of measurement error in radar imaging is target glint. All imaging radars respond to the wavefront which arrives at the antenna. When the target, or some component of the scene being imaged, is a complex scatterer like an aircraft or a refinery pipeyard, the wavefront contains

local ripples which are referred to as glint. The ranging errors caused by glint can be reduced by smoothing, but generally several seconds of averaging are required to obtain a significant error reduction. This time may not be available during a mapping operation.

Another source of error in radar imagery results from multipath reflections from the earth's surface, particularly when target scenes are near the horizon, or when the scenes contain many horizontal reflecting surfaces. These errors show up chiefly along the elevation angle of the imagery. In many instances, these and other sources of image error in active systems can be countered by the use of complex data processing schemes such those as discussed in Chapter Three.

Active MW/MMW imaging systems can scan in a number of different ways. Many systems use the scanning techniques described earlier for passive MW/MMW imagers. Phased arrays are also used to minimize antenna gimbal requirements. In addition, as with laser imaging systems, azimuth sweeping of an elevation-spread fan beam is also used. In the azimuth-swept systems the angular resolution in the elevation direction is limited.

SAR - Another type of imaging system is the Synthetic Aperture Radar (SAR). Synthetic aperture radar is based on the generation of an equivalent large antenna by signal processing means rather than by the actual use of a large antenna. In fact, only a single, relatively small, physical antenna is used in most cases. This system has a number of significant advantages over other types of imaging radars, although the signal processing required is much more complex. The most important advantage is the SAR's ability to improve the azimuth resolution of an airborne ground-mapping radar to a value significantly better than that achievable by use of the radi-

ated beamwidth. For a given antenna aperture, lower-frequency (and less expensive) transmitters and components can be used; the system also takes advantage of the generally better propagation characteristics of the atmosphere at lower frequencies (see Figure 3.4-1b). As a result, SAR imaging is particularly attractive in satellite sensor systems. A specific example (SEASAT) is discussed in Unit Four.

SAR Principle - A long physical antenna is analogous to a long linear array, in which a number of radiating elements are constructed and placed at appropriate points along a straight line. In the use of such a physical linear array, signals are fed simultaneously to each of the elements of the array. Similarly, when it is used as a receiver, the elements receive signals simultaneously; in both the transmitting and receiving modes, waveguide, phase-shifters, or other transmission-line interconnections are used, and interference phenomena are exploited to get a desired effective radiation pattern.

In the synthetic aperture case, usually a single radiating element is used. This antenna is translated (usually horizontally) to take up sequential positions along a line, such as the flight path of the airborne vehicle. At each of these positions, a signal is transmitted and the amplitude and phase of the received signals stored.

After the radiating element has traversed a certain distance, the stored signals approximate very closely the signals that would have been received by the elements of an actual linear array. Consequently, if the signals in storage are subjected to the same operations as those used in forming a physical linear array, the response of a long effective antenna aperture is simulated. This explains the use of the term synthetic aperture.

In the case of an airborne ground-mapping radar system, the antenna usually is mounted to be side-looking, and the motion of the aircraft carries the radiating element to each of the positions of the array. These array positions are the locations of the physical antenna at the times of transmission and reception of the radar signals.

An additional advantage of SAR techniques is that range focusing is possible. A physical linear array can be focused to a specific range. There is then a depth of focus surrounding this range. However, most physical linear arrays are unfocused. This is sometimes expressed by saying that the antenna is focused at infinity. In the SAR, however, it is possible to focus each range separately by the proper adjustment of the phases of the received signals before the summation. The synthetic aperture width is thus made equal to the linear width of the radiated beam at each range. Significantly better angular resolution and signal sensitivity are possible with a focused system, though the signal processing complexity is increased. It is likely that the rapidly advancing state of the art in solid-state microprocessor components will make SAR systems the most widely used airborne MW/MMW imaging systems of the future.

CHAPTER SEVEN

TARGET POSITIONING -- PHOTOGRAMMETRIC TECHNIQUES

The objective of target positioning, in both tactical and strategic applications, is to build, maintain, update, and disseminate lists of potential target locations. The accuracy requirements are both relative and absolute. For example, the location of an ammunition dump for an airstrike may be required with high relative accuracy in relation to local radar reflectors. On the other hand, the location of a hardened ICBM silo must be known with high absolute accuracy in order to provide the necessary capability for interdiction.

A key part of the total target positioning problem is, of course, the initial gathering of intelligence data relative to the areas in question. This subject is beyond the scope of this text and is not discussed here. It is assumed in the following discussion that adequate photographic and other intelligence information is available to the target positioning operators.

3.7.1 Tactical Considerations

Rapid response is the key to providing a target list for tactical operations. This response consists of the following important processes:

- Rapid identification of target location on a stereo pair of photographs [Point Positioning Data Base (PPDB)]
- Accurate measurement of target location relative to known points (control network)

- Translation to required coordinate system for weapon system delivery.

In the first case, the identification of a target's location is primarily a matter of the recognition of distinguishing features on a photograph which mark the site of the proposed target. In a case of a forward observer calling in an air strike, for example, this may be the location of a crossroad. The targeting support system must provide for high-quality optics and a capability for rapid scanning of a PPDB for the area, in order to locate the points in question.

The next part of the targeting problem is for the operator to measure the location of the chosen points relative to a set of <u>control points</u> whose positions have been accurately determined by previous surveys, by location on highly accurate maps, or by some other means. These measurements consist of moving an accurate cursor to the control point position and then measuring, in the coordinates of the PPDB, the distance to the target location. It is very important to recognize that this measurement is three dimensional. The operator must have a stereo view of the area. As a side benefit, stereo viewing tends to enhance the ability of the operator to recognize points of interest.

The control-point to target-point measurements are repeated as many times as necessary, and are made for as many control points as are required to provide enough samples for a statistically accurate estimation of the target position. This process requires some form of computational capability for the operator; typically, a programmable calculator or a microprocessor is included in the mensuration equipment.

The translation to the required coordinates for the particular weapon system can then be computed using the known reference system of the PPDB. This translation process requires computational capability at either the mensuration site or the strike force headquarters.

3.7.2 Strategic Considerations

Targeting for strategic mission planning differs only qualitatively from the processes required for tactical target location. The fundamental differences are in:

- Required response time
- The number and type of target required
- The need for absolute location relative to the earth's surface
- The extreme accuracy required of the measurements.

As for tactical targeting, the strategic process requires an extensive list of easily identified control points whose locations are well specified. However, the areas of the world for which these control points must be defined and the extreme accuracy with which they must be located make the maintenance of a control point file for strategic purposes an extremely taxing operation.

Another aspect of the strategic targeting problem which differs from the tactical problem is the fact that the number of potential targets is very large (on the order of 100,000). Assignment of strike priority and maintenance of the valid target list is a large data processing problem in its own right. Coupled with this is the problem of actually making the measurements for all of the targets with suffi-

cient accuracy to support precise individual strikes as required.

3.7.3 The General Positioning Problem

Given a sufficient stock of background intelligence reports and other data indicating the necessity to add a target to the valid target list, the first prerequisite for determining target location is the availability of the necessary data bases -- that is, collections of accurate stereo pairs of photographs for the given areas. Also needed are supporting digital data bases for use by the computer algorithms which accomplish the final mensuration steps. A major digital data base provides control point identification parameters and location. These are used in concert with imagery or other data to locate targets on the ground.

The development of the control point network includes extrapolation from known target positions to identifiable (but not accurately positioned) features on the earth's surface. It is especially desirable to have a control point network that is sufficiently dense in areas of high interest to support very high accuracy in the mensuration process. At the same time, it is also desirable (because of mathematical constraints in many algorithms) to have a regular grid orientation. Proper selection and measurement of control point data is therefore crucial to the targeting process, whether strategic or tactical.

Some current exploratory development is being directed toward utilization of control point files consisting of photographic or other representations of the neighborhood of control points. In this case, digitized images of the immediate area of the control points or descriptions of the control point

area based on pattern recognition techniques are stored in a digital data base along with the locations of the known control points. These data can then be correlated with digitized photographic images of the area of interest to provide precise location of the control points in the image framework.

UNIT THREE
REVIEW EXERCISES

Chapter One

1. Define the term remote sensing and give several examples.

Chapter Two

2. Briefly summarize the principal characteristics, advantages, and disadvantages of the two major imaging technologies -- photochemical and photo-electronic.

3. It is frequently stated that fluorescent light sources are more efficient than incandescent devices. Explain the meaning of the statement.

4. What would be the color of the light from a hypothetical device that produces light from electricity at close to the theoretical optimum efficiency of 686 lumens/watt?

5. On a very bright day when the sky is clear, the illuminance due to sunlight may be as high as 10^4 footcandle. Express this value in metric units.

6. Although silver fluoride belongs to the class of compounds called silver halides, it is not mentioned in most discussions of silver-based photographic technology. Explain.

7. What is a latent image in photography?

8. Assume that it is necessary to examine an exposed, but undeveloped, strip of aerial film to obtain information about the latent images contained thereon. What are some ways in which this can be done?

9. For the H and D curve shown in Fig. 3.2-5, estimate the value of gamma.

10. Explain the reciprocity law and define the term reciprocity failure.

11. Why is reciprocity failure a more serious problem with color film than with black and white materials?

12. Distinguish between photoconductive and photoemissive devices.

13. As a general rule, there is one characteristic that distinguishes all of the non-silver photographic processes from conventional silver-based technology. Explain.

14. Explain what is meant by the modulation transfer function (MTF).

Chapter Three

15. Distinguish between the concepts of sampling and digitizing.

16. The image enhancement techniques of Chapter Three are described in terms of sampled and digitized images available in digital form (for computer storage and manipulation). These techniques are, therefore, directly applicable to imagery produced by electro-optical devices. How can they be applied to images in continuous form on photographic film?

Chapter Four

17. This exercise refers to Figs. 3.4-1a, b and 3.4-2. For each of the environmental/weather conditions listed below, and for a fixed aperture size, specify which of the atmospheric windows allows ≤ 2 dB/km total one-way attenuation while maximizing the angular resolution obtainable.

    Atmospheric Windows:
    - a. Visible light
    - b. Mid-IR (3-5 μm)
    - c. Far-IR (8-14 μm)
    - d. 220 GHz
    - e. 140 GHz
    - f. 94 GHz
    - g. 35 GHz
    - h. < 20 GHz

    Weather Conditions:
    1. Fog, 400-m visibility
    2. Haze, 1.2-km visibility
    3. Fog, 100-m visibility
    4. Moderate rain (0.4 cm/hr)
    5. Haze, 23-km visibility
    6. Very heavy rain (5 cm/hr)

18. What is meant by Rayleigh scattering? Give an example.

Chapter Five

19. This exercise refers to Figs. 3.5-1 and 3.5-3, and involves two objects:

- Hot metal housing surrounding a tank exhaust vent; temperature, 700 K; IR reflectivity, 0.4

- Black asphalt road; temperature, 270 K; IR reflectivity, 0.05

For each material, find the wavelength at which the peak of its thermal radiation curve occurs. Find the total radiated energy (watt/m^3) at the wavelength of peak emission. Select a detector material that would be most suitable for sensing the radiation from each of the objects.

Chapter Six

20. This exercise refers to a spaceborne millimeter-wave sensor (140 GHz) for passive detection of atmospheric vehicles from radiometric contrast temperature measurements.

    a. Assuming an orbital altitude of 800 km, compute the diameter of the antenna aperture required for resolution of 5-meter target objects.

    b. If the sensor were replaced by one operating at 10 μm, what would be the required antenna aperture for resolution of a 5-meter object?

UNIT THREE
READING LIST

Chapter One

Sabins, F.F., Jr., Remote Sensing: Principles and Interpretation, W.H. Freeman, San Francisco, 1978.

Chapter Two

Thomas, W., Jr., (ed), SPSE Handbook of Photographic Science and Engineering, Wiley-Interscience, New York, 1973.

Arnold, C.R., Rolls, P.J., and Stewart, J.C.C., Applied Photography, Focal Press, 1971.

Jacobson, K.I., and Jacobson, R.E., Imaging Systems: Mechanisms and Applications of Established and New Photosensitive Processes, Wiley, 1976.

Baines, H., The Science of Photography, (3rd ed), Fountain Press, London, 1970.

Biberman, L.M., and Nudelman, S., (eds), Photoelectronic Imaging Devices, Plenum Press, New York, 1971.

Chapter Three

Andrews, H.C., and Hunt, B.R., Digital Image Restoration, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

Rosenfeld, A., and Kak, A.C., Digital Picture Processing, Academic Press, New York, 1976.

Chapter Four

Wallace, J.M., and Hobbs, P.V., Atmospheric Science: An Introductory Survey, Academic Press, New York, 1977.

Zuev, V.E., Propagation of Visible and Infrared Radiation in the Atmosphere, Wiley, New York, 1974.

Chapter Five

Hudson, R.D., Jr., Infrared System Engineering, Wiley-Interscience, New York, 1969.

Wolfe, W.L., and Zissis, G.J., (eds), The Infrared Handbook, Environmental Research Institute of Michigan (ERIM), Ann Arbor, 1978.

Chapter Six

Lloyd, J.M., Thermal Imaging Systems, Plenum Press, New York, 1975.

Brookner, E., Radar Technology, Artech House, Dedham, Mass., 1977.